(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 June 2010 (24.06.2010)

PCT

(10) International Publication Number
WO 2010/069137 A1

(51) International Patent Classification:
C12N 15/82 (2006.01) *A01K 67/00* (2006.01)
A01H 5/00 (2006.01) *C12N 9/16* (2006.01)

(21) International Application Number:
PCT/CN2009/001465

(22) International Filing Date:
16 December 2009 (16.12.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/138,918 18 December 2008 (18.12.2008) US

(71) Applicant *(for all designated States except US)*: **THE UNIVERSITY OF HONG KONG** [CN/CN]; Pokfulam Road, Hong Kong (CN).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **LIM, Boonleong** [CN/CN]; Flat A, Floor 19, Block 1, Phase 1, Residence Bel-Air, 28 Bel-Air Avenue, Island South, Hong Kong (CN).

(74) Agent: **CHINA PATENT AGENT (H.K.) LTD.**; 22/F, Great Eagle Centre, 23 Harbour Road, Wanchai, Hong Kong (CN).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2010/069137 A1

(54) **Title**: METHOD FOR SPEEDING UP PLANT GROWTH AND IMPROVING YIELD BY INTRODUCING PHOSPHATASES IN TRANSGENIC PLANT

(57) **Abstract**: Transgenic plants having increased growth rate, increased sugar content, and increase yield are disclosed, and methods for making the same. The transgenic plants have a gene coding for a phosphatase having a C-terminal motif under control of a heterologous promoter incorporated into the genomic DNA of the plant.

METHOD FOR SPEEDING UP PLANT GROWTH AND IMPROVING YIELD BY INTRODUCING PHOSPHATASES IN TRANSGENIC PLANT

RELATED APPLICATIONS

This application claims priority to provisional application Serial No. 61/138,918, filed on December 18, 2008, which is incorporated herein by reference.

1. TECHNICAL FIELD

The present disclosure provides methods that speeds up plant growth and elevates plant yields by introducing phosphatases with a C-terminal motif into plants. The present disclosure relates to phosphatases with a C-terminal motif, and their respectively encoded protein products, as well as fragments, derivatives, homologues, and variants thereof. Methods for introducing these genes into plants to (1) speed up the growth rate of plants, (2) to increase the sugar contents of plants, and (3) to increase of yield of plants, are provided.

2. BACKGROUND

Purple acid phosphatases (PAPs) catalyze the hydrolysis of a wide range of activated phosphoric acid mono- and di-esters and anhydrides (Klabunde et al., 1996). The PAP proteins are characterized by seven conserved amino acid residues (shown in bold face) in the five conserved motifs X**D**XX, X**D**XX**Y**, GNH(D/E), XXX**H**, X**H**X**H**, which are involved in the coordination of the dimetal nuclear center (Fe^{3+}-Me^{2+}) in the active site (Li et al., 2002), where Me is a transition metal and Me^{2+} is mostly found to be Fe^{2+} in mammalian, and Zn^{2+}, or Mn^{2+} in plants (Klabunde and Krebs, 1997; Schenk *et al.*, 1999).

Purple acid phosphatases are distinguished from the other phosphatases by their characteristic purple color, which is caused by a charge transfer transition at 560nm from a metal-coordinating tyrosine to the metal ligand Fe^{3+} (Klabunde and Krebs, 1997; Schenk *et al.*, 2000). Different from the other acid phosphatases, PAPs are insensitive to inhibition by tartrate, so they are also known as tartrate-resistant acid phosphatases (TRAPs).

The biochemical properties of some plant PAPs have been characterized, firstly in red kidney bean, and later in soybean suspension cell, soybean seedlings, rice culture cells, spinach leaves, sweet potato tubers, tomato, yellow lupin seeds,

medicago and Arabidopsis, etc. (Schenk *et al.*, 1999). Plant PAPs are generally considered to mediate phosphorus acquisition and redistribution based on their ability to hydrolyze phosphate compounds (Cashikar *et al.*, 1997; Bozzo *et al.*, 2004; Lung *et al.*, 2008). Regulation of some plant PAPs transcripts by external phosphate level in medium or soil, strongly suggest their involving in phosphate acquisition. For example, the transcription level of Medicago MtPAP1 in roots was increased under P stress, implicating a role in P acquisition or internal mobilization (Xiao *et al.*, 2005; Xiao *et al.*, 2006). Some plant PAPs could be secreted from root cells to extracellular environment, then hydrolyze various phosphate esters. Lung et al. purified a secreted PAP phosphatase from tobacco, which could hydrolyze broad substrates and help to alleviate P starvation (Lung *et al.*, 2008). Certain plant PAPs can also hydrolyze phytate, a major storage compound of phosphorus in plants. Hegeman and Grabau (2001) purified a novel PAPs (GmPhy) from the cotyledon of the germinating soybean seedlings. GmPhy was introduced into soybean tissue culture and was assayed to show phosphatase activity. Most recently, AtPAP15 and 23 in Arabidopsis sharing high sequence homology (73-52%) with this soybean PAP, were found to exhibit phytase activity (Zhu *et al.*, 2005; Zhang *et al.*, 2008).

Besides involvement in P acquisition, plant PAPs may perform some other physiological roles. For example, the PAPs AtACP5 (AtPAP17), SAP1, and SAP2 (del Pozo et al., 1999; Bozzo et al., 2002) display not only phosphatase but also peroxidase activity, suggesting their involvement in the removal of reactive oxygen compounds in plant organs. A pollen-specific PAP from Ester lily was suggested to function as an iron carrier in mature pollen (Kim and Gynheung, 1996). Other studies indicate that plant PAPs may also be involved in NaCl stress adaption or cell regeneration (Kaida, 2003; Liao *et al.*, 2003).

In the Arabidopsis genome, twenty-nine potential PAP genes were identified based on sequence comparison. Twenty-four of these putative enzymes contain seven conserved amino-acids residues involved in metal binding. One (AtPAP13) lacked four of these seven residues, and the other four (AtPAP14, 16, 28 and 29) lacked either the first, the second, or both motifs of the five conserved motifs. Twenty-eight are actively transcribed in Arabidopsis (Zhu *et al.*, 2005).

To date, relatively little is known about AtPAPs biochemical properties and physiological roles, though several members have been characterized (del Pozo *et al.*, 1999). AtPAP17 (AtACP5) was first known to be induced by phosphorus starvation.

The transcription of AtPAP17 was also responsive to ABA, salt stress (NaCl), oxidative stress (H_2O_2) and leaves senescence, according to GUS activity assay. No alteration in the expression of AtPAP17 was observed during the nitrogen or potassium starvation, and paraquat or salicylic acid. Like the other type5 acid phosphatases, AtPAP17 displayed peroxidation activity, which may be involved in the metabolism of reactive oxygen species in stressed or senescent parts of plants.

Besides AtPAP17, several AtPAPs were found to be involved in phosphorus metabolism in Arabidopsis. Root secretion of AtPAP12 was induced by P stress, and its regulation was mainly at transcriptional level (Patel *et al.*, 1998; Coello, 2002/11). AtPAP4, as well as AtPAP10, AtPAP11 and AtPAP12 were involved in phosphorus starvation response since their transcription levels increased during phosphate deprivation (Li *et al.*, 2002; Wu *et al.*, 2003). In contrast, AtPAP20, 21 and 22 were irrespective to P starvation and expressed constitutively in Pi sufficient or deficient condition. Fluorescent signals were detected in the cytoplasm via the baculovirus expression system, indicating that they may function in the cytoplasm (Li and Wang, 2003).

AtPAP26 was purified and characterized from Pi-starved Arabidopsis suspension cell culture (Veljanovski *et al.*, 2006). It exists as a homodimer with 55 kDa glycosylated protein, showing wide substrate specificity with the highest activity against phosphoenolpyruvate (PEP) and polypeptide phosphate. AtPAP26 also displayed alkaline peroxidase activity with the probable roles in the metabolism of reactive oxygen species. Proteomic study suggested that it may be localized in vacuole, and involved in recycling Pi from intracellular P metabolites (Shimaoka *et al.*, 2004).

PAPs can act on a wide range of substrates, but not all of them exhibit phytase activity. An enzyme assay involving the GST-AtPAP23 fusion protein revealed that AtPAP23 exhibits phytase activity. A GUS study showed that AtPAP23 is exclusively expressed in the flower of the Arabidopsis, and may play certain roles in flower development (Zhu *et al.*, 2005). In a recent report, a recombinant AtPAP15 expressed and partial purified in *E. coli* and yeast was also found to exhibit phytase activity (Zhang *et al.*, 2008). It was proposed that AtPAP15 may be involved in ascorbic acid biosynthesis with the end product *myo*-inositol of phytate hydrolysis as the precursor of ascorbic acid synthesis.

As stated above, most of the functions of characterized plant PAPs are related to phosphorus metabolism. None of the functionally or biochemically characterized plant PAPs carry transmembrane motif, and none of them were shown to be associated with membrane. Furthermore, to date, no AtPAPs or any plant PAPs, have been showed to affect sugar signalling and carbon metabolism in plant.

The first report of transgenic expression of plant PAP in plant was reported in 2005 (Xiao *et al.*, 2005). The PAP-phosphatase gene from Medicago (*MtPHY1*) was expressed in transgenic Arabidopsis, resulting in increased capacity of P acquisition from phytate in agar culture (Xiao *et al.*, 2005). Nonetheless, the growth performance of the plants was not reported to be different under normal growth.

3. SUMMARY

The present disclosure provides a method that speeds up plant growth and elevates plant yields by introducing phosphatases with a C-terminal motif into plants. Phosphatases with a C-terminal motif, and their respectively encoded protein products, as well as fragments, derivatives, homologues, and variants thereof are disclosed. Methods for introducing this class of genes into plants to speed up the growth rate of plants, to increase the sugar contents of plants, and to increase of yield of plants, are provided. Without wishing to be bound by any particular theory, the C-terminal motif is believed to function as a transmembrane structural element (transmembrane motif).

As stated above in the Background section, most of the functions of characterized plant PAPs are related to phosphorus metabolism. None of the functionally or biochemically characterized plant PAPs carry transmembrane motif, and none of them were shown to be associated with membrane. Furthermore, to date, no AtPAPs or any plant PAPs, have been showed to affect sugar signalling and carbon metabolism in plant.

The first report of transgenic expression of plant PAP in plant was reported in 2005 (Xiao *et al.*, 2005). The PAP-phosphatase gene from Medicago (*MtPHY1*) was expressed in transgenic Arabidopsis, resulting in increased capacity of P acquisition from phytate in agar culture. Nonetheless, the growth performance of the plants was not reported to be different under normal growth.

We also produced transgenic tobacco and Arabidopsis that overexpressed AtPAP15, a PAP with phosphatase activity, which does not carry any C-terminal motif equivalent to that of AtPAP2; phosphatase activity was secreted into

extracellular growth medium. Significant secretion of phosphatase activity was observed in the transgenic plants and the transgenic plants showed larger biomass than the control plants in agar and soil supplemented with exogenous phytate. Higher P content was also obtained in overexpressed transgenic lines in phytate treatment. However, the growth of transgenic plants overexpressing AtPAP15 did not show any difference in growth phenotypes when it was compared with the wild-type, under treatments of K-P or No-P, or in soil.

Here, we have developed a technology to speed up plant growth and improve seed yield by overexpressing a phosphatase with a C-terminal motif in plants. An example is the use of a purple acid phosphatase (PAP). This disclosure is the first report to show that overexpressing a phosphatase with a C-terminal motif in transgenic plant is able to speed up the growth of the plants, to increase the sugar contents of plants, and to increase the yield of plants, by altering the carbon metabolism of the plants.

The present advances are based, in part, on the characterization of a group of purple acid phosphatases (SEQ ID NOS: 1-8 and 18-47) from plants and the observations that overexpression of a purple acid phosphatase (AtPAP2, SEQ ID NO:1) of this group in plants resulted in rapid plant growth, higher sugar content, and higher yield. Accordingly, nucleotide sequences of a group of purple acid phosphatase genes (SEQ ID NOs:1, 3, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 and 46), which share a C-terminal motif/domain, from plants and amino acid sequences of their encoded proteins (SEQ ID NOS:2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47), as well as fragments, derivatives, homologues, and variants thereof, as defined herein, are disclosed. Furthermore, nucleic acid molecules encoding the polypeptides of interest,, and include cDNA, genomic DNA, and RNA, are disclosed.

As used herein, italicizing the name of a gene shall indicate the gene, in contrast to its encoded protein or polypeptide product which is indicated by the name of the gene in the absence of any italicizing. For example, "*Gene*" shall mean the *Gene* gene, whereas "Gene" shall indicate the protein or polypeptide product of the *Gene* gene.

In one embodiment, isolated nucleic acid molecules hybridize under stringent conditions, as defined herein, to nucleic acids having the sequence of SEQ ID NOS: 1, 3, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 or 46, or homologues

thereof, wherein the nucleic acid molecules encode proteins or polypeptides which exhibit at least one structural and/or functional feature of the polypeptides of the invention.

Another embodiment includes, nucleic acid molecules, which are suitable for use as primers or hybridization probes for the detection of nucleic acids encoding one of the disclosed phosphatase polypeptides or other sequences.

Yet another embodiment includes vectors, *e.g.*, recombinant expression vectors, comprising a nucleic acid molecule of the invention. Furthermore, host cells containing such a vector or engineered to contain and/or express a nucleic acid molecule of the invention and host cells containing a nucleotide sequence of the invention operably linked to a heterologous promoter are disclosed.

A further embodiment includes methods for preparing a polypeptide of the invention by a recombinant DNA technology in which the host cells containing a recombinant expression vector encoding a polypeptide of the invention or a nucleotide sequence encoding a polypeptide of the invention operably linked to a heterologous promoter, are cultured, and the polypeptide of the invention are produced.

In still furhter another embodiment, a transgenic plant contains a nucleic acid molecule which encodes an isolated polypeptides or proteins comprising the five conserved motifs of purple acid phosphatases, including X**D**XX, X**D**XX**Y**, GNH(D/E), XXX**H**, X**H**X**H**, and linked to a C-terminal motif.

Embodiments further provide antibodies that immunospecifically bind a polypeptide of the invention. Such antibodies include, but are not limited to, antibodies from various animals, humanized, chimeric, polyclonal, monoclonal, bi-specific, multi-specific, single chain antibodies, Fab fragments, $F(ab')_2$ fragments, disulfide-linked Fvs, fragments containing either a VL or VH domain or even a complementary determining region (CDR), that immunospecifically binds to a polypeptide of the invention.

In an additional embodiment, method for detecting the presence, activity or expression of a polypeptide of the invention or similar polypeptide in a biological material, such as cells, culture media, and so forth are provided. The increased or decreased activity or expression of the polypeptide in a sample relative to a control sample can be determined by contacting the biological material with an agent that can detect directly or indirectly the presence, activity or expression of the polypeptide of

the invention. In a particular embodiment, such an agent is an antibody or a fragment thereof which immunospecifically binds to a one of the disclosed polypeptides.

In a still another embodiment, a fusion protein comprising a bioactive molecule and one or more domains of a disclosed polypeptide or fragment thereof is provided. In particular, fusion proteins comprising a bioactive molecule recombinantly fused or chemically conjugated (including both covalent and non-covalent conjugations) to one or more domains of a disclosed polypeptide or fragments thereof.

We also produced transgenic tobacco and Arabidopsis that overexpressed AtPAP15, a PAP with phosphatase activity, which does not carry any C-terminal motif and was found to be secreted into extracellular growth medium. Significant secretion of phosphatase activity was observed in the transgenic plants and the transgenic plants showed larger biomass than the control plants in agar and soil supplemented with exogenous phytate. Higher P content was also obtained in overexpressed transgenic lines in phytate treatment. However, the growth of transgenic plants overexpressing AtPAP15 did not show any difference in growth phenotypes when it was compared with the wild-type, under treatments of K-P or No-P, or in soil.

In conclusion, this disclosure is the first report to show that overexpressing a phosphatase with a C-terminal motif in transgenic plant is able to speed up the growth of the plants, to increase the sugar contents of plants, and to increase the yield of plants, by altering the carbon metabolism of the plants.

3.1 Definitions

The term "acidic" or "acid pH" as used herein refers to a pH value of less than about 6.0.

The term "homologue" as used herein refers to a polypeptide that possesses a similar or identical function to polypeptides encoded by SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47, and/or a fragment of these polypeptides, that do not have an identical amino acid sequence of these polypeptides and/or a fragment of these polypeptides. A polypeptide that has a similar amino acid sequence included in the definition of the term "homologue" includes a polypeptide that satisfied at least one of the following: (i) polypeptide having an amino acid

sequence that is one or more of at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, and at least about 98% identical. (ii) a polypeptide encoded by a nucleotide sequence that is one or more of at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, and at least about 98%identical and/or conservatively substituted to one or more of the nucleotide sequences encoding the polypeptides of SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47, and/or a fragment of the these polypeptides; (iii) a polypeptide encoded by a nucleotide sequence that hybridizes under stringent conditions as defined herein to one or more of nucleotide sequences SEQ ID NOS: 1, 3, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 and 6; (iv) a polypeptide having an amino acid sequence that is one or more of at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, and at least about 98% identical and/or conservatively substituted; (v) a nucleic acid sequence encoding an amino acid sequence that is one or more of at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, and at least about 98% identical and/or conservatively substituted; (vi) a fragment of any of the polypeptides or nucleic acid sequences described in (i) through (v) having one of at least 20 amino acid residues, at least 25 amino acid residues, at least 40 amino acid residues, at least 80 amino acid residues, at least 90 amino acid residues, at least 100 amino acid residues, at least 125 amino acid residues, at least 150 amino acid residues, at least 175 amino acid residues, at least 200 amino acid residues, at least 225 amino acid residues, at least 250 amino acid residues, at least 275 amino acid residues, at least 300 amino acid residues, at least 325 amino acid residues, at least 350 amino acid residues, or at least 375 amino acid residues; (vii) a polypeptide with similar structure and function or a nucleotide sequence encoding a polypeptide with similar structure and function, exhibiting the antigenicity, immunogenicity, catalytic activity, and other readily assayable activities, to polypeptides encoded by SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47, and/or a fragment of these polypeptides, refers to a polypeptide that has a similar secondary, tertiary, or quaternary structure of these polypeptides, or a fragment of these polypeptides. The structure of a polypeptide can be determined by methods known to those skilled in the art, including but not limited to, X-ray crystallography, nuclear magnetic resonance,

and crystallographic electron microscopy. The term "homologue" is used herein to describe a sequence that has sequence homology. A sequence having sequence homology can be made using standard molecular biology techniques including site-directed mutagenesis including insertion or deletion of sequences. The term "homologue" is not limited to homologous genes or proteins originating from different species and expressly includes artificial modification to the sequences disclosed herein.

The term "conservatively substituted variant" refers to a polypeptide or a nucleic acid sequence encoding a homologue polypeptide in which one or more amino acid residues or codons have been modified by conservative substitution with an amino acid residue or a codon coding for an amino acid residue of similar chemical-type, as described below.

The term "an antibody or an antibody fragment which immunospecifically binds to polypeptides encoded by SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47, " as used herein refers to an antibody or a fragment thereof that immunospecifically binds to polypeptides encoded by SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47, or a fragment of these polypeptide and does not non-specifically bind to other polypeptides. An antibody or a fragment thereof that immunospecifically binds to polypeptides encoded by SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47, or a fragment of these polypeptide, may cross-react with other antigens. Preferably, an antibody or a fragment thereof that immunospecifically binds to polypeptides encoded by SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47, or a fragment of these polypeptides, does not cross-react with other antigens. An antibody or a fragment thereof that immunospecifically binds to polypeptides encoded by SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47, or a fragment of these polypeptide, can be identified by, for example, immunoassays or other techniques known to those skilled in the art. An antibody or an antibody fragment which immunospecifically binds polypeptides encoded by SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47,may be interchangeably referred to as "anti-PAP antibody".

The term "derivative" as used herein refers to a given peptide or protein that is otherwise modified, *e.g.*, by covalent attachment of any type of molecule, preferably having bioactivity, to the peptide or protein, including the incorporation of non-

naturally occurring amino acids. The resulting bioactivity retains one or more biological activities of the peptide protein.

The term "fragment" as used herein refers to a fragment of a nucleic acid molecule containing one of at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 100, at least about 150, at least about 200, at least about 250, at least about 300, at least about 350, at least about 400, at least about 450, at least about 500, at least about 550, at least about 600, at least about 650, at least about 700, at least about 750, at least about 800, at least about 850, at least about 900, at least about 950, at least about 1000, at least about 1050, at least about 1100, at least about 1150, at least about 1200, at least about 1250, at least about 1300, at least about 1350, from about 500 to about 2000, from about 1000 to about 2000 from about 200 to about 500, from about 500 to about 1000, form about 1000 to about 1500, and from about 1500 to about 2000nucleic acid bases in length of the relevant nucleic acid molecule and having at least one functional feature of the nucleic acid molecule (or the encoded protein has one functional feature of the protein encoded by the nucleic acid molecule); or a fragment of a protein or a polypeptide containing one or more of at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 90, at least about 100, at least about 120, at least about 140, at least about 160, at least about 180, at least about 200, at least about 220, at least about 240, at least about 260, at least about 280, at least about 300, at least about 320, at least about 340, at least about 360, from about 250 to about 660, from about 350 to about 660, form about 450 to about 660, and form about 550 to about 660 amino acid residues in length of the relevant protein or polypeptide and having at least one functional feature of the protein or polypeptide, such functional features include ability to bind a Fe^{3+}-Me^{2+} dimetal nuclear center and form a C-terminal motif .

An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid molecule. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular materials, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized, but excludes nucleic acid molecules present

in recombinant DNA libraries. In a preferred embodiment, nucleic acid molecules encoding the disclosed polypeptides/proteins are isolated or purified.

The term "operably linked" as used herein refers to when transcription under the control of the "operably linked" promoter produces a functional messenger RNA, translation of which results in the production of the polypeptide encoded by the DNA operably linked to the promoter.

The term "under stringent condition" refers to hybridization and washing conditions under which nucleotide sequences having homology to each other remain hybridized to each other. Such hybridization conditions are described in, for example but not limited to, *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6.; *Basic Methods in Molecular Biology*, Elsevier Science Publishing Co., Inc., N.Y. (1986), pp. 75-78, and 84-87; and *Molecular Cloning*, Cold Spring Harbor Laboratory, N.Y. (1982), pp. 387-389, and are well known to those skilled in the art. A preferred, non-limiting example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC), 0.5% SDS at about 68°C followed by one or more washes in 2X SSC, 0.5% SDS at room temperature. Another preferred, non-limiting example of stringent hybridization conditions is hybridization in 6X SSC at about 45°C followed by one or more washes in 0.2X SSC, 0.1% SDS at about 50-65°C.

The term "variant" as used herein refers either to a naturally occurring allelic variation of a given peptide or a recombinantly prepared variation of a given peptide or protein in which one or more amino acid residues have been modified by amino acid substitution, addition, or deletion.

The term “aligned” as used herein refers to a homology alignment between two or more sequences using a standard algorithm such as BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

The term “predicted to form a transmembrane motif by TMHMM analysis” or “predicted to form a C-terminal motif by TMHMM analysis” (http://www.cbs.dtu.dk/services/TMHMM/) herein refers to a probability that is equal to or greater than about 0.5.

BRIEF DESCRIPTION OF THE FIGURES

The following figures illustrate the embodiments and are not meant to limit the scope of the invention encompassed by the claims.

FIG. 1 shows the phylogenetic tree of PAP-like sequences in the Arabidopsis genome. Twenty-nine PAPs were aligned using *ClustalX* and the phylogenetic tree was created by the neighbor-joining algorithm of the MEGA4 program. The accession numbers of the PAP-like, transmembrane-like C-terminal motif containing, polypeptide from *Zea mays* (ZmPAP2) and *Oryza sativa* (OsPAP2) were ACG47621 and BAC15853.1, respectively.

FIG 2A is the amino acid alignment of AtPAP2 with other PAP sequences, showing the full length of each sequence. These sequences include homologous sequences from *B. napus* (BnPAP2), *G. max* (GmPAP2) and *Z. may* (ZmPAP2). The five conserved motifs (X**D**XX, X**D**XX**Y**, G**N**H(D/E), XXX**H**, X**H**X**H**) are boxed. Residues in shades have low or no homology. Hydrophobic motifs at the C-termini of these polypeptides are underlined by a bar (614^{th}-636^{th} amino acid), which is absent from the sequence of AtPAP15. As shown, AtPAP15 does not have a C-terminal region corresponding to the other PAP sequences.

FIG. 2B is the amino acid alignment of the C-terminal transmembrane-like motifs in AtPAP2 and its homologous sequences.

FIG. 3 shows that a unique hydrophobic motif is present at the C- termini of AtPAP2 and ZmPAP2 by TMHMM analysis. This transmembrane-like C-terminal motif is absent from AtPAP15.

FIG. 4 shows the characteristics of the T-DNA lines. The T-DNA line (Salk_013567) was obtained from TAIR. The *AtPAP2* genomic sequence carries two exons and the T-DNA was inserted in exon 2 and causes a disruption of the *AtPAP2* mRNA (a). Three PCR primers (A, B and C) were designed for the differentiation of the wild-type (WT) and the T-DNA line (*atpap2-8*) and they were used for PCR screening of genomic DNA extracted from WT and the T-DNA line (b). Total RNA was extracted from 10-day-old seedlings grown on MS with 2% sucrose using the TRIzol RNA isolation method and were used for RT-PCR (c). 50μg of seedlings proteins were loaded for Western blotting studies, using the anti-AtPAP2 specific antiserum (Section 6.3).

FIG. 5 is the schematic diagram of the expression vector pBV-AtPAP2. CaMV 35S: 35S promoter of the cauliflower mosaic virus; NOS: polyadenylation signal of nopaline synthase gene; *aadA*: bacterial streptomycin/spectinomycin resistance gene encoding aminoglycoside-3"-adenyltransferase; pNOS:BAR: bialaphos resistance gene under the control of the nopaline synthase promoter; bom: basis of mobility from pBR322; ColE1: replication origin from pBR322; pVS1-REP: replication origin from pVS1; pVS1-STA: STA region from pVS1 plasmid; LB: left border T-DNA repeat; RB: right border T-DNA repeat. (Hajdukiewicz *et al.*, 1994).

FIG. 6A shows the results of the Western blot analysis of the overexpression lines (OE), wild-type (WT), T-DNA and the complementation lines (CP) of AtPAP2 and FIG. 6B shows the results of the Western blot analysis of the overexpression lines (C-15) and wild-type (WT) of AtPAP15.

FIG. 7 shows the expression analysis of AtPAP2. The mRNA expression profile was analysed by the Spot History program of NASC (a). The protein expression profiles of 30 day old, soil-grown plant (b), seedlings germinated on MS agar (c) and 2 week old plants transferred to Pi-sufficient/Pi-deficient MS agar for 3 days (d), were analyzed by Western blotting using the anti-PAP2 antiserum.

FIG. 8 shows the growth performance of the wild-type, T-DNA and overexpression lines in soil. Seeds were germinated in MS agar with 2% sucrose for 10 days. Seedlings with 2 small visible rosette leaves (~ 1 mm) were transferred to soil and grown under 16h/8h light/dark cycles.

FIG. 9 shows the levels of sucrose and glucose in the rosette leaves of 21-day-old, soil grown seedlings.

FIG. 10 shows the recovery of various lines after prolonged darkness treatment. Seeds were germinated in MS agar with 2% sucrose for 10 days. Seedlings with 2 small visible rosette leaves (~ 1 mm) were transferred to soil and grown for 12 days under 16h/8h light/dark cycles. The lights of the growth chamber were then switched off for 12 days and the plants were allowed to recover under 16h/8h light/dark cycles for 1 week. n = 9-12 per line.

FIG. 11. Detection of AtPAP2 protein in subcellular fractions by Western blotting. Mito.: Mitochondria; Chlorop.: Chloroplasts.

Fig. 12. shows a schematic representation of two vector constructs incorporating the AtPAP2 gene.

Fig. 13. shows Western blot analysis results for overexpression of AtPAP2 proteins missing the C-terminal motif.

DETAILED DESCRIPTION

5.1 Method of speeding up plant growth and improving crop yield

The present disclosure provides a method that speeds up plant growth and elevates plant yields by introducing phosphatases with a C-terminal motif into plants. In an embodiment, the present disclosure relates to a class of genes of purple acid phosphatases, and their respectively encoded protein products, as well as fragments, derivatives, homologues, and variants thereof. Methods for introducing this class of genes into plants to speed up the growth rate of plants, to increase the sugar contents of plants, and to increase of yield of plants, are provided.

A group of purple acid phosphatases (PAPs) which carry seven conserved amino acid residues (shown in bold face) in the five conserved motifs X**D**XX (example GDXG (SEQ ID NO: 48)), X**D**XX**Y** (SEQ ID NO: 49), G**N**H(D/E) (SEQ ID NOS: 50-51), XXX**H** (example ZXGH (SEQ ID NO: 52)), X**H**X**H** (SEQ ID NO: 53), where X is any amino acid and Z is any amino acid selected from L, I, V, F, andM, and a transmembrane-like motif at their C-termini were identified in the genomes of a number of plants (FIGS. 1, 2A, and 2B). The presence of the C-terminal transmembrane-like motif enables the localization of this group of PAPs to the membrane fraction (FIGS. 3 and 11). This property makes this group of PAPs differ from the other previously characterized PAPs because all previously characterized PAPs did not carry any C-terminal motif (FIGS. 2A, 2B, and 3). By using the protein sequence of a representative gene of this group, AtPAP2, to blast the NCBI database and various EST databases, a number of genomic or cDNA sequences were identified to encodes polypeptides that carry the five conserved motifs X**D**XX, X**D**XX**Y**, G**N**H(D/E), XXX**H**, X**H**X**H** of PAPs and a transmembrane motif at their C-termini (FIG. 2A).

The introduction of a representative gene of this group of phosphatases, *AtPAP2*, into the genome of Arabidopsis by transgenic technology produced transgenic Arabidopsis that grew faster than the wild-type plants (FIG. 8), and the yield of seeds were elevated by approximately 40% (Table 3). However, transgenic

plant that expressed AtPAP15 did not show these phenotypes. The sugar contents, including glucose and sucrose, in the leaf of the transgenic lines, were also found to be higher than that of the wild-types (FIG. 9).

Thus, this disclosure provides a method that speeds up plant growth and elevates plant yields by introducing phosphatases into plants. In an embodiment, a group of genes of purple acid phosphatases, and their respectively encoded protein products, as well as fragments, derivatives, homologues, and variants thereof are described.

5.2 Homologues, Derivatives, and Variants of Phosphatases

In addition to the nucleic acid molecules (SEQ ID NOS: 1, 3, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 and 46) and polypeptides (SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47) described in claims 9-16, the nucleic acid molecules and polypeptides also encompass those nucleic acid molecules and polypeptides having a common biological activity, similar or identical structural domain and/or having sufficient nucleotide sequence or amino acid identity (homologues) to those of the nucleic acid molecules and polypeptides described above.

Such common biological activities of the polypeptides include antigenicity, immunogenicity, catalytic activity especially phosphatase activity, ability to bind a Fe^{3+}-Me^{2+} dimetal nuclear center, fold into or form a transmembrane-like C-terminal motif and other activities readily assayable by the skilled artisan.

A polypeptide that has a similar amino acid sequence (homologue) refers to a polypeptide that satisfied at least one of the following: (i) a polypeptide having an amino acid sequence that is one of at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, and at least about 95%, and at least about 98% identical and/or conservatively substituted to the amino acid sequence of a AtPAP2 (SEQ ID NO: 2) and/or other PAPs with a transmembrane-like C-terminal motif including SEQ ID NOS: 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and/or 47, a fragment of AtPAP2, and having at least one biological feature of the described polypeptides; (ii) a polypeptide encoded by a nucleotide sequence that is one of at least about 30%, at least about 40%, at least about 40%, at least about 50%, at least about 60%, at least

about 70%, at least about 80%, at least about 90%, at least about 95%, and at least about 98% identical to the nucleotide sequence encoding *AtPAP2* (SEQ ID NO: 1) and/or other PAPs with a transmembrane-like C-terminal motif including SEQ ID NOS: 3, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 and/or 46, a fragment of AtPAP2 and having at least one structural and/or biological feature of AtPAP2; (iii) a polypeptide encoded by a nucleotide sequence that hybridizes under stringent conditions as defined herein to a nucleotide sequence encoding *AtPAP2* (SEQ ID NO: 1) and/or other PAPs with a motif including SEQ ID NOS: 3, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 and/or 46, a fragment of AtPAP2 and having at least one structural and/or biological feature of AtPAP2. A polypeptide with similar structure to AtPAP2, or a fragment of AtPAP2, refers to a polypeptide that has a similar secondary, tertiary, or quaternary structure of AtPAP2, a fragment of AtPAP2 and has at least one functional feature of a AtPAP2, including one or more of ability to bind a Fe^{3+}-Me^{2+} dimetal nuclear center and fold into or form a transmembrane-like C-terminal motif. The structure of a polypeptide can be determined by methods known to those skilled in the art, including but not limited to, X-ray crystallography, nuclear magnetic resonance, and crystallographic electron microscopy.

Those having skill in the art will readily recognized that mutations, deletions or insertions can be made in any of the sequences disclosed herein, including SEQ ID NOS: 1-8 and 18-47, without affecting function. Sequences useful in practicing the embodiments include sequences having homology to SEQ ID NOS: 1-8 and 18-47 and being a protein, polypeptide, or polynucleotide coding for such protein or peptide having functionality to bind a dimetal nuclear center (Fe^{3+}-Me^{2+}) and being a protein, polypeptide, or polynucleotide coding for such protein or peptide having a C-terminal motif. That is, those skilled in the art will recognize that many mutations can be made to any of SEQ ID NOS: 1-8 and 18-47 without affecting the catalytic functionality nor interrupting the transmembrane-like C-terminal motif. Such modified sequences that maintain catalytic activityand a transmembrane-like C-terminal motif are defined as homologues to SEQ ID NOS: 1-8 and 18-47 and are including within the scope of useful sequences.

In one embodiment, such homologues can have about 30% or more identity to the sequences disclosed herein. In another embodiment, such homologues can have about 40% or more identity to the sequences disclosed herein. In yet another

embodiment, such homologues can have about 50% or more identity to the sequences disclosed herein. In sill yet another embodiment, such homologues can have about 60% or more identity to the sequences disclosed herein. In even sill yet another embodiment, such homologues can have about 70% or more identity to the sequences disclosed herein. In a further embodiment, such homologues can have about 80% or more identity to the sequences disclosed herein. In yet a still further embodiment, homologues can have about 90% or more identity to the sequences disclosed herein. In a still further embodiment, homologues can have about 98% or more identity to the sequences disclosed herein.

Those having skill in the art will recognize that mutations can be made to proteins and peptides and/or to polynucleotides coding for protein and peptides or complementary thereto that substitute amino acid residue for other amino acids residues having similar chemical properties (conservative substitutions) and that such mutations are less likely to cause structural changes that affect functionality including catalytic acitivty and/or the function of a transmembrane-like C-terminal motif. Conservatively substituting amino acids are substituting an amino acid residue belong to any of the following 11 chemical groups with another amino acid from the same chemical group: (1) acidic (negatively charged) amino acids such as aspartic acid and glutamic acid; (2) basic (positively charged) amino acids such as arginine, histidine, and lysine; (3) neutral polar amino acids such as glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; (4) neutral nonpolar (hydrophobic) amino acids such as alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; (5) amino acids having aliphatic side chains such as glycine, alanine, valine, leucine, and isoleucine; (6) amino acids having aliphatic-hydroxyl side chains such as serine and threonine; (7) amino acids having amide-containing side chains such as asparagine and glutamine; (8) amino acids having aromatic side chains such as phenylalanine, tyrosine, and tryptophan; (9) amino acids having basic side chains such as lysine, arginine, and histidine; (10) amino acids having sulfur-containing side chains such as cysteine and methionine; (11); amino acids having similar geometry and hydrogen bonding patterns such as aspartic acid, asparagine, glutamic acid and glutamine.

In one embodiment, homologues can have about 30% or more identity and/or conservative substitutions to the sequences disclosed herein. In another embodiment, homologues can have about 40% or more identity and/or conservative substitutions to

the sequences disclosed herein. In yet another embodiment, homologues can have about 50% or more identity and/or conservative substitutions to the sequences disclosed herein. In still yet another embodiment, homologues can have about 60% or more identity and/or conservative substitutions to the sequences disclosed herein. In a further embodiment, homologues can have about 70% or more identity and/or conservative substitutions to the sequences disclosed herein. In a still further embodiment, homologues can have about 80% or more identity and/or conservative substitutions to the sequences disclosed herein. In still another embodiment, homologues can have about 90% or more identity and/or conservative substitutions to the sequences disclosed herein. In still another further embodiment, homologues can have about 98% or more identity and/or conservative substitutions to the sequences disclosed herein.

Embodiments further provide isolated nucleic acid molecules which comprise or consist of one or more of at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 100, at least about 150, at least about 200, at least about 250, at least about 300, at least about 350, at least about 400, at least about 450, at least about 500, at least about 550, at least about 600, at least about 650, at least about 700, at least about 750, at least about 800, at least about 850, at least about 900, at least about 950, at least about 1000, at least about 1050, at least about 1100, at least about 1150, at least about 1200, at least about 1250, at least about 1300, at least about 1350, from about 500 to about 2000, from about 1000 to about 2000, from about 200 to about 500, from about 500 to about 1000, form about 1000 to about 1500, and from about 1500 to about 2000 nucleotides of the nucleotide sequences of SEQ ID NOS: 1, 3, 5,7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 and 46, or a complement thereof encoding a protein or polypeptide having one or more activity of the amino acid sequences of their encoded proteins (SEQ ID NOS: 2, 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47). The activity includes one or more of antigenicity, immunogenicity, catalytic activity (*e.g.*, phosphatase activity), ability to bind a Fe^{3+}-Me^{2+} dimetal nuclear center, fold into or form a transmembrane-like C-terminal motif, and other activities readily assayable.

Embodiments provide isolated polypeptides or proteins consisting of an amino acid sequence that contains one of about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65,

at least about 70, at least about 75, at least about 80, at least about 90, at least about 100, at least about 120, at least about 140, at least about 160, at least about 180, at least about 200, at least about 220, at least about 240, at least about 260, at least about 280, at least about 300, at least about 320, at least about 340, at least about 360, from about 250 to about 660, from about 350 to about 660, form about 450 to about 660, and form about 550 to about 660 amino acid bases in length of the relevant protein or polypeptide and having at least one functional feature of the protein or polypeptide, such functional features including ability to bind a Fe^{3+}-Me^{2+} dimetal nuclear center and form a transmembrane-like C-terminal motif.

Additional embodiments are any of the phosphatases and homologues thereof with the identity and/or conservative substitutions to SEQ ID NOS: 1-8 and 18-47 described above that additionally consist of a protein, polypeptide, or polynucleotide encoding a protein having the five conserved motifs in purple acid phosphatases, including X**D**XX, X**D**XX**Y**, GNH(D/E), XXX**H**, X**H**X**H,** where X is any amino acid. In one embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding one of the sequences YHVCIGNHEYDF (SEQ ID NO: 54) and YHVCIGNHEYDW (SEQ ID NO: 55). In one embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having one of the sequences YHVCIGNHEYD(W/F) (SEQ ID NO: 54) and YHVCIGNHEYN(W/F) (SEQ ID NO: 55) or a protein, polypeptide, or polynucleotide encoding a homologue to one of the foregoing sequences with only conservative substitutions, as described above, to those sequences. In yet another embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having one of the sequences GNHE (SEQ ID NO: 51) and GNHD (SEQ ID NO: 50). In still yet another embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having one of the sequences GNHE (SEQ ID NO: 51) and GNHD (SEQ ID NO: 50) or a protein, polypeptide, or polynucleotide encoding a protein having a homologous sequence to one of the foregoing sequences SEQ IN NOS: 50-51 with only conservative substitutions,.

Additional embodiments are any of the phosphatases and homologues thereof with the identity and/or conservative substitutions to SEQ ID NOS: 1-8 and 18-47

described above that additionally consist of a protein, polypeptide, or polynucleotide encoding a sequence having at least about 70% or more identity and/or conservative substitutions to amino acid residues 302-315 of SEQ ID NO: 2 when such sequence is aligned with SEQ ID NO: 2. In another embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a protein having about 80% or more identity and/or conservative substitutions to amino acid residues 302-315 of SEQ ID NO: 2 when such sequence is aligned with SEQ ID NO: 2. In another embodiment, the described phosphatases and homologues consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having at least about 70% or more identity to the sequence HIGDISYARGYSW (SEQ ID NO: 56). In another embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having the sequence HIGDISYARGYSW (SEQ ID NO: 56) or a protein, polypeptide, or polynucleotide encoding a protein having a homologous sequence to the foregoing sequences with only conservative substitutions, as described above, to those sequences.

In another embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having at least about 70% or more identity to the sequences KEKLTVSFVGNHDGEVHD (SEQ ID NO: 57), KERLTLSYVGNHDGEVHD (SEQ ID NO: 58), REKLTLTYVGNHDGQVHD (SEQ ID NO: 59), and KEKLTLTYIGNHDGQVHD (SEQ ID NO: 60). In still yet another embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having one or more of the sequences KEKLTVSFVGNHDGEVHD (SEQ ID NO: 57), KERLTLSYVGNHDGEVHD (SEQ ID NO: 58), REKLTLTYVGNHDGQVHD (SEQ ID NO: 59), and KEKLTLTYIGNHDGQVHD (SEQ ID NO: 60) or a protein, polypeptide, or polynucleotide encoding a protein having a homologous sequence to one of the foregoing sequences with only conservative substitutions, as described above, to those sequences.

In a further embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having the sequence (F/Y)(V/I)GNHDGXXH (SEQ ID NOS: 61-64), where the first residue of the sequence can be F or Y and the second residue of the

sequence can be V or I. In a still further embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having the sequence (F/Y)(V/I)GNHDGXXH (SEQ ID NOS: 61-64), where the first residue of the sequence can be F or Y and the second residue of the sequence can be V or I, or a protein, polypeptide, or polynucleotide encoding a protein having a homologous sequence to the foregoing sequence with only conservative substitutions, as described above, to the foregoing sequence. In a yet still further embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having the sequence (F/Y)(V/I)GNHDGXXH (SEQ ID NOS: 61-64), where the first residue of the sequence can be F or Y and the second residue of the sequence can be V or I, or a protein, polypeptide, or polynucleotide encoding a protein having a homologous sequence having at least about 70% identity and/or conservative substitution, as described above, to the foregoing sequence.

In one embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having at least about 60% or more identity and/or conservative substitutions to amino acid residues 614-636 of SEQ ID NO: 2 (SEQ ID NO: 65) and/or having at least about 60% or more identity and/or conservative substitutions to the sequence of 23 amino acid residues of SEQ ID NOS: 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33 and 47 aligned with residues 614-636 of SEQ ID NO: 2 (SEQ ID NO: 65) , and where amino acid residues aligned with amino acid residues 614-636 of SEQ ID NO: 2 are predicted to form a transmembrane-like C-terminal motif by TMHMM analysis (http://www.cbs.dtu.dk/services/TMHMM/). In one embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having at least about 70% or more identity and/or conservative substitutions to amino acid residues 614-636 of SEQ ID NO: 2 (SEQ ID NO: 65) and/or having at least about 60% or more identity and/or conservative substitutions to the sequence of 23 amino acid residues of SEQ ID NOS: 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33 and 47 aligned with residues 614-636 of SEQ ID NO: 2 (SEQ ID NO: 65) , and where amino acid residues aligned with amino acid residues 614-636 of SEQ ID NO: 2 (SEQ ID NO: 65) are predicted to form a transmembrane-like C-terminal motif by TMHMM analysis (http://www.cbs.dtu.dk/services/TMHMM/). In one embodiment, the described

phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having at least about 80% or more identity and/or conservative substitutions to amino acid residues 614-636 of SEQ ID NO: 2 and/or having at least about 60% or more identity and/or conservative substitutions to the sequence of 23 amino acid residues of SEQ ID NOS: 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33 and 47 aligned with residues 614-636 of SEQ ID NO: 2 (SEQ ID NO: 65) , and where amino acid residues aligned with amino acid residues 614-636 of SEQ ID NO: 2 (SEQ ID NO: 65) are predicted to form a transmembrane-like C-terminal motif by TMHMM analysis (http://www.cbs.dtu.dk/services/TMHMM/). In one embodiment, the described phosphatases and homologues thereof consist of a protein, polypeptide, or polynucleotide encoding a sequence of amino acid residues having at least about 90% or more identity and/or conservative substitutions to amino acid residues 614-636 of SEQ ID NO: 2 and/or having at least about 90% or more identity and/or conservative substitutions to the sequence of 23 amino acid residues of SEQ ID NOS: 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33 and 47 aligned with residues 614-636 of SEQ ID NO: 2, and where amino acid residues aligned with amino acid residues 614-636 of SEQ ID NO: 2 are predicted to form a transmembrane-like C-terminal motif by TMHMM analysis (http://www.cbs.dtu.dk/services/TMHMM/).

In one embodiment, the described phosphatases or phosphatase genes consist of a protein, polypeptide, or polynucleotide encoding the sequence (L/M/V)-(L/M/V)-Z-(G/A)-(V/A/L)-Z-Z-G-(F/Y)-X-Z-G (SEQ ID NO: 66), where Z is any of the hydrophobic residues L, I, V, F, and M. In another embodiment, the described phosphatase or phosphatase genes consist of a protein, polypeptide, or polynucleotide encoding the sequence (L/M/V)-(L/M/V)-Z-(G/A)-(V/A/L)-Z-Z-G-(F/Y)-X-Z-G (SEQ ID NO: 66), or a protein, polypeptide, or polynucleotide encoding a sequence having at least 70% identity and/or conservative substitution to the foregoing sequence.

Embodiments also encompass derivatives of the disclosed polypeptides. For example, but not by way of limitation, derivatives may include peptides or proteins that have been modified, *e.g.*, by glycosylation, acetylation, pegylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques, including,

but not limited to, specific chemical cleavage, acetylation, formylation, etc. Additionally, the derivative may contain one or more non-classical amino acids.

In another aspect, an isolated nucleic acid molecule encodes a variant of a polypeptide in which the amino acid sequences have been modified by genetic engineering so that biological activities of the polypeptides are either enhanced or reduced, or the local structures thereof are changed without significantly altering the biological activities. In one aspect, these variants can act as either agonists or as antagonists. An agonist can retain substantially the same or a portion of the biological activities of the polypeptides and an antagonist can inhibit one or more of the activities of the polypeptides. Such modifications include amino acid substitution, deletion, and/or insertion. Amino acid modifications can be made by any method known in the art and various methods are available to and routine for those skilled in the art.

For example, mutagenesis may be performed in accordance with any of the techniques known in the art including, but not limited to, synthesizing an oligonucleotide having one or more modifications within the sequence of a given polypeptide to be modified. Site-specific mutagenesis can be conducted using specific oligonucleotide sequences which encode the nucleotide sequence containing the desired mutations in addition to a sufficient number of adjacent nucleotides in the polypeptide. Such oligonucleotides can serve as primers which can form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 15 to about 75 nucleotides or more in length is preferred, with about 10 to about 25 or more residues on both sides of the junction of the sequence being altered. A number of such primers introducing a variety of different mutations at one or more positions can be used to generate a library of mutants.

The technique of site-specific mutagenesis is well known in the art, as described in various publications (*e.g.*, Kunkel *et al.*, Methods Enzymol., 154:367-82, 1987, which is hereby incorporated by reference in its entirety). In general, site-directed mutagenesis is performed by first obtaining a single-stranded vector or melting apart of two strands of a double stranded vector which includes within its sequence a DNA sequence which encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as T7 DNA polymerase, in order to complete the

synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform or transfect appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement. As will be appreciated, the technique typically employs a phage vector which exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phages are readily commercially available and their use is generally well known to those skilled in the art. Double stranded plasmids are also routinely employed in site directed mutagenesis which eliminates the step of transferring the gene of interest from a plasmid to a phage.

Alternatively, the use of PCR with commercially available thermostable enzymes such as *Taq* DNA polymerase may be used to incorporate a mutagenic oligonucleotide primer into an amplified DNA fragment that can then be cloned into an appropriate cloning or expression vector. See, *e.g.*, Tomic *et al.*, *Nucleic Acids Res.*, 18(6):1656, 1987, and Upender *et al.*, *Biotechniques*, 18(1):29-30, 32, 1995, for PCR-mediated mutagenesis procedures, which are hereby incorporated in their entireties. PCR employing a thermostable ligase in addition to a thermostable polymerase may also be used to incorporate a phosphorylated mutagenic oligonucleotide into an amplified DNA fragment that may then be cloned into an appropriate cloning or expression vector (see *e.g.*, Michael, *Biotechniques*, 16(3):410-2, 1994, which is hereby incorporated by reference in its entirety).

Other methods known to those skilled in art of producing sequence variants of a given polypeptide or a fragment thereof can be used. For example, recombinant vectors encoding the amino acid sequence of the polypeptide or a fragment thereof may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants.

Optionally, the amino acid residues to be modified are surface exposed residues. Additionally, in making amino acid substitutions, preferably the amino acid residue to be substituted is a conservative amino acid substitution, for example, a polar residue is substituted with a polar residue, a hydrophilic residue with a hydrophilic residue, hydrophobic residue with a hydrophobic residue, a positively charged residue with a positively charged residue, or a negatively charged residue

with a negatively charged residue. Moreover, the amino acid residue that can be modified is not highly or completely conserved across strains or species and/or is critical to maintain the biological activities of the protein.

Accordingly, included in the scope of the disclosure are nucleic acid molecules encoding a polypeptide of the invention that contains amino acid modifications that are not critical to its biological activity.

5.3 Fusion Proteins

The present disclosure further encompasses fusion proteins in which the polypeptides or fragments thereof, are recombinantly fused or chemically conjugated (*e.g.*, covalent and non-covalent conjugations) to heterologous polypeptides (*i.e.*, an unrelated polypeptide or portion thereof, preferably at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90 or at least 100 amino acids of the polypeptide) to generate fusion proteins. The fusion can be direct, but may occur through linker sequences.

In one aspect, the fusion protein comprises a polypeptide which is fused to a heterologous signal sequence at its N-terminus. For example, the signal sequence naturally found in the polypeptide can be replaced by a signal sequence which is derived from a heterologous origin. Various signal sequences are commercially available.

In another embodiment, a polypeptide can be fused to tag sequences, *e.g.*, a hexa-histidine peptide, among others, many of which are commercially available. As described in Gentz *et al.*, 1989, *Proc. Natl. Acad. Sci. USA*, 86:821-824, for instance, hexa-histidine provides for convenient purification of the fusion protein. Other examples of peptide tags are the hemagglutinin "HA" tag, which corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson *et al.*, 1984, *Cell*, 37:767) and the "flag" tag (Knappik *et al.*, 1994, *Biotechniques*, 17(4):754-761). These tags are especially useful for purification of recombinantly produced polypeptides.

Fusion proteins can be produced by standard recombinant DNA techniques or by protein synthetic techniques, *e.g.*, by use of a DNA synthesizer. For example, a nucleic acid molecule encoding a fusion protein can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification

of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, *e.g.*, *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, 1992).

The nucleotide sequence coding for a fusion protein can be inserted into an appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted protein-coding sequence.

In a specific embodiment, the expression of a fusion protein is regulated by an inducible promoter.

5.4 Preparation of Transgenic Plants

Carbon flow is a key process in plant biology and high energy carbon molecules (e.g. glucose) were harvested by plant through photosynthesis. The carbon molecules were then converted into more complicated carbohydrate molecules such as starch, cellulose, etc. Cellulose is the major component of cell wall and starch is the major storage form of glucose in plant cells and plant seeds. Therefore, the efficiency and/or the equilibrium of the carbon flow process become a limiting factor for plant growth and crop yield.

The present disclosure is based upon the discovery that overexpression of a membrane-bound phosphatase can enhance the growth performance of plants by altering its carbon metabolism, as indicated by, for example, a faster growth rate, a higher sugar contents, and a higher seed yield.

In an embodiment, the present disclosure provides a transgenic plant containing a nucleic acid molecule that encodes and expresses a phosphatase having a C-terminal transmembrane-like domain. The transgenic plants disclosed herein have faster growth rate, and higher seed yield to comparable unengineered plants *i.e.* same species (strain). In a specific embodiment, such a phosphatase is from a plant species having a phosphatase activity and a C-terminal motif. In another embodiment, a transgenic plant disclosed herein comprises a nucleic acid molecule encoding phosphatase and expresses AtPAP2 (SEQ ID NO: 2) and/or other PAPs with a C-terminal motif including one or more of SEQ ID NOS: 4, 6, 8, 19, 21, 23, 25, 27, 29,

31, 33, 35, 37, 39, 41, 43, 45 and 47. In another embodiment, the phosphatase is expressed on cellular membrane, for example, the ER or the Golgi apparatus. Such a membrane expression of a phosphatase in plants can be achieved by fusing onto the C-terminus with a nucleotide sequence encoding a C-terminal motif peptide which can efficiently attach the phosphatase upon translation thereof from the cells of a given plant. Accordingly, in another embodiment, a transgenic plant comprises a nucleic acid molecule encoding phosphatase and expresses AtPAP2 (SEQ ID NO: 2) and/or other PAPs with a C-terminal motif including SEQ ID NOS: 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and 47, except that all or a portion, particularly an N-terminal portion, of amino acid residues 1 to 80, preferably all or a portion of amino acid residues 1 to 30, of SEQ ID NO: 2 or all or a portion, particularly an N-terminal portion, of amino acid residues 1 to 80, preferably all or a portion of amino acid residues 1 to 30, of SEQ ID NOS: 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 or 47, are replaced by a heterologous plant signal peptide by genetic engineering. In such a transgenic plant, the phosphatases are directed to various organelles/compartments of the cells. In another embodiment, a transgenic plant comprises a nucleic acid molecule encoding phosphatase and expresses homologues, derivatives, and/or fragments thereof having at least one functional feature and/or structural feature of a phosphatase polypeptide. In all embodiments where all or a portion of the N-terminal portion of SEQ ID NOS: 4, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 and/or 47 are replaced, the embodiments include homologues to such sequences, as described above, having at least one functional feature and/or structural feature of a phosphatase polypeptide. In yet another embodiment, a transgenic plant comprises a nucleic acid molecule that hybridizes under stringent conditions, as defined herein, to a nucleic acid molecule having the sequence of SEQ ID NOS: 1, 3, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 or 46, or a complement thereof, and encodes a protein or polypeptide that exhibits at least one structural and/or functional feature of the disclosed phosphatase polypeptides. Specifically, the production of transgenic plant that overexpressed a membrane-bound phosphatase, which contributes to improving plant physiology, such as plant growth rate and characteristics, for example, in seed yield, is provided.

Accordingly, also provided are chimeric gene constructs for genetic modification of plants to increase their growth rate and improve the yield. The

chimeric gene constructs comprise a sequence that encodes substantially solely for a phosphatase enzyme that carry a C-terminal transmembrane-like motif. Such a phosphatase enzyme can be derived from the purple acid phosphatase family. In a specific embodiment, the chimeric gene constructs comprise a nucleic acid having the sequence of SEQ ID NOS: 1 , 3, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 or 46. In another embodiment, the chimeric gene constructs comprise a nucleic acid molecule that encodes a homologue or fragment thereof having at least one functional feature and/or structural feature of a phosphatase polypeptide. In another specific embodiment, the chimeric gene constructs comprise a sequence that hybridizes under stringent conditions, as defined herein, to a nucleic acid having the sequence of SEQ ID NOS: 1, 3, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 or 46, or a complement thereof, wherein the sequence encodes a protein or a polypeptide that exhibits at least one structural and/or functional feature of the phosphatase polypeptides. Furthermore, the phosphatases encoded by the nucleic acid molecules contained in the chimeric gene constructs can be any other phosphatases that have similar structural characteristics, such as having a C-terminal transmembrane-like motif, to those of the phosphatases described herein. Such phosphatase include, but not limited to, the following polypeptides: Purple acid phosphatases from *Zea mays* (Accession No: ACG47621); and *Oryza sativa* (Accession No: BAC15853.1).

The phosphatase-coding sequence is operatively linked to upstream and downstream regulatory components, preferably heterologous to the phosphatase sequence; for example CMV 35S promoter, which acts to cause expression of the gene (production of the enzyme) in plant cells (*see* Section 6.2). When a construct containing a gene for a phosphatase according to this disclosure, is introduced into plant cells by a conventional transformation method, such as microparticle bombardment, Agrobacterium infection, or microinjection, the gene is expressed in the cells under the control of the regulatory sequences. The expressed phosphatase successfully interacts with the biosynthetic machinery that is naturally present in the plant cells to alter the carbon metabolism. By altering the carbon metabolism, the method described herein also favors the growth rate of the plant, resulting in faster growth rate and higher yield. Thus, the time required for the maturation of the plant and the time required for flowering is shortened. Also provided are methods for

increasing growth rate and yield of plants, comprising the step of inserting into such plant cells or the cells of such whole plants a chimeric gene construct.

In specific embodiments, Arabidopsis (see Section 6) was adopted as the model system. An overexpression construct the gene coding for phosphatase were introduced into Arabidopsis.

In an embodiment, the phosphatase from *Arabidopsis* is used. The results obtained with this disclosure indicate that the growth rate and the seed yield of transgenic Arabidopsis were enhanced by overexpressing this gene (*see* Section 6.5 and FIG. 8 and Table 3).

While any plant species can be modified using the expression cassette and methods described herein, preferably included without limitation are species from the following genera with representative species in parentheses:

Monocots: genera Asparagus (asparagus), Bromus (cheatgrass), Hemerocallis (daylily), Hordeum (barley), Lolium (ryegrass), Oryza (rice), Panicum (Switchgrass), Pennisetum (fountaingrass), Saccharum (Sugar cane), Sorghum, Trigonella (fenu grass), Triticum (wheat), Zea (corn); and

Dicots: genera Antirrhinum (flower sp.), Arabidopsis (thaliana), Arachis (peanut), Atropa (deadly nightshade), Brassica (rapeseed), Browallia, Capsicum (pepper), Carthamus (safflower), Cichorium (chicory), Citrus (orange, lemon), Chrysanthemum, Cucumis (cucumber), Datura (thorn apple), Daucus (carrot), Digitalis (foxglove), Fragaria (strawberry), Geranium (flower sp.), Glycine (soybean), Helianthus (sunflower), Hyscyamus, Ipomoea (morning glory), Latuca (lettuce), Linum (linseed), Lotus (flower sp.), Lycopersicon (tomato), Majorana, Malva (cotton), Manihot, Medicago (alfalfa), Nemesia, Nicotiana (tobacco), Onobrychis, Pelargonium (citrosa), Petunia (flower sp.), Ranunculus (flower sp.), Raphanus (radishes), Salpiglossis, Senecio (flower sp.), Sinapis (albae semen), Solanum (potato), Trifolium (clovers), Vigna (mungbean, faba bean), Vitis (grape).

Genetic engineering of plants can be achieved in several ways. The most common method is *Agrobacterium*-mediated transformation. In this method, *A. tumefaciens*, which in nature infects plants by inserting tumor causing genes into a plant's genome, is altered. Selected genes are engineered into the T-DNA of the bacterial Ti (tumor-inducing) plasmid of A. tumefaciens in laboratory conditions so that they become integrated into the plant chromosomes when the T-DNA is transferred to the plant by the bacteria's own internal transfer mechanisms. The only

essential parts of the T-DNA are its two small (25 base pair) border repeats, at least one of which is needed for plant transformation. The bacterial genes encoding for plant hormones that promote tumor growth are excised from the T-DNA and replaced with a sequence of DNA that typically contains: a selectable marker (*e.g.* an antibiotic-resistance gene; usually kanamycin resistance), a restriction site - a site with a specific sequence of nucleotides where a restriction enzyme will cut the DNA, and the desired genes to be incorporated into the plant (B. Tinland, 1996. The integration of T-DNA into plant genomes. Trends in Plant Science 1,178-184; D. Grierson (ed.) 1991. Plant Genetic Engineering. Blackie, Glasgow). *Agrobacterium* can be added to plant protoplasts (plant cells with cell walls removed) in culture, that are then allowed to regenerate cell walls at which point non-transformed plants are killed with antibiotics for which the transformed plants have been given resistance genes. Plantlets are then regenerated from the surviving transformed cells using standard plant tissue culture techniques. In an alternative technique, sterile disks or fragments of vegetative portions of plants are place in liquid culture medium with *Agrobacterium*, then hormones are used to induce rooting thereby regenerate plantlets which are grown on selection media. A third technique for delivering genes is possible for some plants such as *Arabidopsis* where the *Agrobacterium* or even "naked" DNA can be infused through the seed coat to cause transformation (Clough SJ and Bent AF, 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J 16:735-43).

The biolistic method for genetic engineering of plants was developed more recently and is becoming more widely employed. In this method, very small particles (microprojectiles) of tungsten or gold coated with biologically active DNA are propelled at high-velocities into plant cells using an electrostatic pulse, air pressure, or gunpowder percussion. As the particles pass through the cell, the DNA dissolves and can then integrate into the genome of that cell and its progeny. It has been demonstrated this method can produce stable transformants (Christou, P., *et al.*, 1988. Stable transformation of soybean callus by DNA-coated gold particles, *Plant Physiology* 87:671-674). The method can be practiced on whole plants and is particularly effective on meristematic tissue. It is also capable of delivering DNA either to the nucleus or into mitochondria (Johnston, S.A., *et al.*, 1988. Mitochondrial transformation in yeast by bombardment with microprojectiles (Science 240,1538-41)

and chloroplasts (Svab, Z., *et al.*, 1990, Stable transformation of plastids in higher plants, *Proc Natl Acad Sci.* USA 87, 8526-8530).

The electroporation method of plant genetic engineering has met with less success. In this technique, protoplasts in culture take up pure DNA when treated with certain membrane-active agents or with electroporation, a rapid pulse of high-voltage direct current. Once the DNA has entered the protoplast it can be integrated into the cells genome. Standard tissue culture techniques are then used to regenerate transgenic plants.

The microinjection method of plant genetic engineering is perhaps the most difficult. In this method, DNA is microinjected into target plant cells using very thin glass needles in a method similar to that used with animals. The technique is laborious, ineffective, and impractical for generating large numbers of transgenic plants.

The method chosen for genetically engineering plants is most often dependent on the targeted plant species and which methods have been proven effective therein.

5.5 Preparation of Antibodies

Antibodies which specifically recognize one of the described phosphatase polypeptides or fragments thereof can be used for detecting, screening, and isolating the polypeptide of the invention or fragments thereof, or similar sequences that might encode similar enzymes from the other organisms. For example, in one specific embodiment, an antibody which immunospecifically binds AtPAP2 or fragments thereof can be used for various *in vitro* detection assays, including enzyme-linked immunosorbent assays (ELISA), radioimmunoassays, Western blot, etc., for the detection of the polypeptide of the invention or fragments, derivatives, homologues, or variants thereof, or similar molecules having the similar enzymatic activities as the phosphatase polypeptides, in samples, for example, a biological material, including plant cells, plants, food, drinks, or any materials derived from plants.

Antibodies specific for the described phosphatase polypeptides can be generated by any suitable method known in the art. Polyclonal antibodies to an antigen-of-interest can be produced by various procedures well known in the art. For example, an antigen derived from the phosphatase polypeptide can be administered to various host animals including, but not limited to, rabbits, mice, rats, etc., to induce

the production of antisera containing polyclonal antibodies specific for the antigen. Various adjuvants may be used to increase the immunological response, depending on the host species, and include but are not limited to, Freund's (complete and incomplete) adjuvant, mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful adjuvants for humans such as BCG (*Bacille Calmette-Guerin*) and *Corynebacterium parvum*. Such adjuvants are also well known in the art.

Monoclonal antibodies can be prepared using a wide variety of techniques known in the art including the use of hybridoma, recombinant, and phage display technologies, or a combination thereof. For example, monoclonal antibodies can be produced using hybridoma techniques including those known in the art and taught, for example, in Harlow *et al.*, *Antibodies: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, *et al.*, in: *Monoclonal Antibodies and T-Cell Hybridomas*, pp. 563-681 (Elsevier, N.Y., 1981) (both of which are incorporated by reference in their entireties). The term "monoclonal antibody" as used herein is not limited to antibodies produced through hybridoma technology. The term "monoclonal antibody" refers to an antibody that is derived from a single clone, including any eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced.

Methods for producing and screening for specific antibodies using hybridoma technology are routine and well known in the art. In a non-limiting example, mice can be immunized with an antigen of interest or a cell expressing such an antigen. Once an immune response is detected, *e.g.*, antibodies specific for the antigen are detected in the mouse serum, the mouse spleen is harvested and splenocytes isolated. The splenocytes are then fused by well known techniques to any suitable myeloma cells. Hybridomas are selected and cloned by limiting dilution. The hybridoma clones are then assayed by methods known in the art for cells that secrete antibodies capable of binding the antigen. Ascites fluid, which generally contains high levels of antibodies, can be generated by inoculating mice intraperitoneally with positive hybridoma clones.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, Fab and $F(ab')_2$ fragments may be produced by proteolytic cleavage of immunoglobulin molecules, using enzymes such as papain (to

produce Fab fragments) or pepsin (to produce $F(ab')_2$ fragments). $F(ab')_2$ fragments contain the complete light chain, and the variable region, the CH1 region and the hinge region of the heavy chain.

The antibodies or fragments thereof can be also produced by any method known in the art for the synthesis of antibodies, in particular, by chemical synthesis or preferably, by recombinant expression techniques.

The nucleotide sequence encoding an antibody may be obtained from any information available to those skilled in the art (*i.e.*, from Genbank, the literature, or by routine cloning). If a clone containing a nucleic acid encoding a particular antibody or an epitope-binding fragment thereof is not available, but the sequence of the antibody molecule or epitope-binding fragment thereof is known, a nucleic acid encoding the immunoglobulin may be chemically synthesized or obtained from a suitable source (*e.g.*, an antibody cDNA library, or a cDNA library generated from, or nucleic acid, preferably poly A+ RNA, isolated from any tissue or cells expressing the antibody, such as hybridoma cells selected to express an antibody) by PCR amplification using synthetic primers hybridizable to the 3' and 5' ends of the sequence or by cloning using an oligonucleotide probe specific for the particular gene sequence to identify, *e.g.*, a cDNA clone from a cDNA library that encodes the antibody. Amplified nucleic acids generated by PCR may then be cloned into replicable cloning vectors using any method well known in the art.

Once the nucleotide sequence of the antibody is determined, the nucleotide sequence of the antibody may be manipulated using methods well known in the art for the manipulation of nucleotide sequences, *e.g.*, recombinant DNA techniques, site directed mutagenesis, PCR, etc. (see, for example, the techniques described in Sambrook *et al.*, *supra*; and Ausubel *et al.*, eds., 1998, *Current Protocols in Molecular Biology*, John Wiley & Sons, NY, which are both incorporated by reference herein in their entireties), to generate antibodies having a different amino acid sequence by, for example, introducing amino acid substitutions, deletions, and/or insertions into the epitope-binding domain regions of the antibodies or any portion of antibodies which may enhance or reduce biological activities of the antibodies.

Recombinant expression of an antibody requires construction of an expression vector containing a nucleotide sequence that encodes the antibody. Once a nucleotide sequence encoding an antibody molecule or a heavy or light chain of an antibody, or portion thereof has been obtained, the vector for the production of the antibody

molecule may be produced by recombinant DNA technology using techniques well known in the art as discussed in the previous sections. Methods which are well known to those skilled in the art can be used to construct expression vectors containing antibody coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. The nucleotide sequence encoding the heavy-chain variable region, light-chain variable region, both the heavy-chain and light-chain variable regions, an epitope-binding fragment of the heavy- and/or light-chain variable region, or one or more complementarity determining regions (CDRs) of an antibody may be cloned into such a vector for expression. Thus-prepared expression vector can be then introduced into appropriate host cells for the expression of the antibody. Accordingly, embodiments include host cells containing a polynucleotide encoding an antibody specific for the disclosed phosphatase polypeptides or fragments thereof.

The host cell can be co-transfected with two expression vectors, the first vector encoding a heavy chain derived polypeptide and the second vector encoding a light chain derived polypeptide. The two vectors may contain identical selectable markers which enable equal expression of heavy and light chain polypeptides or different selectable markers to ensure maintenance of both plasmids. Alternatively, a single vector may be used which encodes, and is capable of expressing, both heavy and light chain polypeptides. In such situations, the light chain should be placed before the heavy chain to avoid an excess of toxic free heavy chain (Proudfoot, 1986, *Nature*, 322:52; and Kohler, 1980, *Proc. Natl. Acad. Sci. USA*, 77:2197). The coding sequences for the heavy and light chains may comprise cDNA or genomic DNA.

In another embodiment, antibodies can also be generated using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the polynucleotide sequences encoding them. In a particular embodiment, such phage can be utilized to display antigen binding domains, such as Fab and Fv or disulfide-bond stabilized Fvs, expressed from a repertoire or combinatorial antibody library (*e.g.*, human or murine). Phage expressing an antigen binding domain that binds the antigen of interest can be selected or identified with antigen, *e.g.*, using labeled antigen or antigen bound or captured to a solid surface or bead. Phages used in these methods are typically filamentous phage, including fd and M13. The antigen binding

domains are expressed as a recombinantly fused protein to either the phage gene III or gene VIII protein. Examples of phage display methods that can be used to make the immunoglobulins, or fragments thereof, include those disclosed in Brinkman *et al.*, 1995, *J. Immunol. Methods* 182:41-50; Ames *et al.*, 1995, *J. Immunol. Methods* 184:177-186; Kettleborough *et al.*, 1994, *Eur. J. Immunol.*, 24:952-958; Persic *et al.*, 1997, *Gene*, 187:9-18; Burton *et al.*, 1994, *Advances in Immunology* 57:191-280; PCT application No. PCT/GB91/01134; PCT publications WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Patent Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

As described in the above documents, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired fragments, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, *e.g.*, as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and $F(ab')_2$ fragments can also be employed using methods known in the art such as those disclosed in PCT publication WO 92/22324; Mullinax *et al.*, 1992, *BioTechniques* 12(6):864-869; and Sawai *et al.*, 1995, AJRI 34:26-34; and Better *et al.*, *Science*, 240:1041-1043, 1988 (each of which is incorporated by reference in its entirety). Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Patent Nos. 4,946,778 and 5,258,498; Huston *et al.*, 1991, *Methods in Enzymology* 203:46-88; Shu *et al.*, 1993, PNAS 90:7995-7999; and Skerra *et al.*, 1988, *Science* 240:1038-1040.

Once an antibody molecule has been produced by any methods described above, it may then be purified by any method known in the art for purification of an immunoglobulin molecule, for example, by chromatography (*e.g.*, ion exchange, affinity, particularly by affinity for the specific antigen after Protein A or Protein G purification, and sizing column chromatography), centrifugation, differential solubility, or by any other standard techniques for the purification of proteins. Further, the antibodies or fragments thereof may be fused to heterologous polypeptide sequences described herein or otherwise known in the art to facilitate purification.

Antibodies fused or conjugated to heterologous polypeptides may be used in *in vitro* immunoassays and in purification methods (*e.g.*, affinity chromatography)

well known in the art. See *e.g.*, PCT publication Number WO 93/21232; EP 439,095; Naramura *et al.*, 1994, *Immunol. Lett.* 39:91-99; U.S. Patent 5,474,981; Gillies *et al.*, 1992, PNAS 89:1428-1432; and Fell *et al.*, 1991, *J. Immunol.* 146:2446-2452, which are incorporated herein by reference in their entireties.

Antibodies may also be attached to solid supports, which are particularly useful for immunoassays or purification of the described polypeptides or fragments, derivatives, homologues, or variants thereof, or similar molecules having the similar enzymatic activities as the polypeptide of the invention. Such solid supports include, but are not limited to, glass, cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene.

5.6 Detection Assays

An exemplary method for detecting the presence or absence of an over-expressed phosphatase polypeptide or an inserted phosphatase-encoding nucleic acid in a biological sample involves obtaining a biological sample from various sources and contacting the sample with a compound or an agent capable of detecting a polypeptide or nucleic acid (*e.g.*, mRNA, genomic DNA) such that the presence of a heterologous polypeptide or nucleic acid is detected in the sample. An exemplary agent for detecting mRNA or genomic DNA encoding an inserted phosphatase polypeptide is a labeled nucleic acid probe capable of hybridizing to mRNA or genomic DNA encoding any of the described phosphatase polypeptides. The nucleic acid probe can be, for example, a full-length cDNA, such as the nucleic acid of SEQ ID NOS: 1 ,3, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 or 46, or a portion thereof, such as an oligonucleotide of at least one of at least about 15, at least about 20, at least about 25, at least about 30, at least about 50, at least about 100, at least about 250, at least about 500, or more nucleotides in length and sufficient to specifically hybridize under stringent conditions to a mRNA or genomic DNA encoding a polypeptide of the invention.

An exemplary agent for detecting an over-expressed phosphatase polypeptide is an antibody capable of binding to a phosphatase polypeptide product of an inserted phosphatase gene, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof

(*e.g.*, Fab or $F(ab')_2$) can be used. See also the detailed descriptions about antibodies in Section 5.5.

The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The detection method can be used to detect mRNA, protein, or genomic DNA in a sample *in vitro* as well as *in vivo*. For example, in vitro techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of a heterologous polypeptide include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of a heterologous polypeptide include introducing into a subject organism a labeled antibody directed against the polypeptide. For example, the antibody can be labeled with a radioactive marker whose presence and location in the subject organism can be detected by standard imaging techniques, including autoradiography.

In a specific embodiment, the methods further involve obtaining a control sample from a control subject, contacting the control sample with a compound or agent capable of detecting an over-expressed polypeptide product or the mRNA transcription product or genomic DNA encoding an inserted phospatase gene, such that the presence of the polypeptide or mRNA or genomic DNA encoding the phosphatase polypeptide is detected in the sample, and comparing the presence of the phosphatase polypeptide or mRNA or genomic DNA encoding the polypeptide in the control sample with the presence of the polypeptide or mRNA or genomic DNA encoding endogenous phosphatase polypeptides in the test sample.

Embodiments also encompasse kits for detecting the presence of a heterologous polypeptide or nucleic acid in a test sample.

The kit, for example, can comprise a labeled compound or agent capable of detecting the polypeptide or mRNA encoding the polypeptide in a test sample and means for determining the amount of the polypeptide or mRNA in the sample (*e.g.*, an

antibody which binds the polypeptide or an oligonucleotide probe which binds to DNA or mRNA encoding the polypeptide). Kits can also optionally include instructions for use.

For antibody-based kits, the kit can comprise, for example: (1) a first antibody (*e.g.*, attached to a solid support) which binds to a phosphatase polypeptide; and, optionally, (2) a second, different antibody which binds to either the polypeptide or the first antibody and is conjugated to a detectable agent.

For oligonucleotide-based kits, the kit can comprise, for example: (1) an oligonucleotide, *e.g.*, a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding an inserted phosphatase polypeptide or (2) a pair of primers useful for amplifying a nucleic acid molecule encoding an inserted phosphatase polypeptide. The kit can also comprise, *e.g.*, a buffering agent, a preservative, or a protein stabilizing agent. The kit can also comprise components necessary for detecting the detectable agent (*e.g.*, an enzyme or a substrate). The kit can also contain a control sample or a series of control samples which can be assayed and compared to the test sample contained. Each component of the kit is usually enclosed within an individual container and all of the various containers are within a single package along with instructions for use.

5.7 Commercial Application of Transgenic Plants

The transgenic plants generated can have many useful applications, including food, feed, biomass, biofuels (starch, cellulose, seed lipids) and wood pulp. The enhanced growth rate of the transgenic plants may provide additional carbon dioxide fixation per hectare of land per year and thus generate carbon credits.

6. EXAMPLES

The following examples illustrate the cloning of AtPAP2, its overexpression in transgenic Arabidopsis, and the characterization of the transgenic plants. These examples should not be construed as limiting. The following examples illustrate some embodiments. Unless otherwise indicated in the following examples and elsewhere in the specification and claims, all parts and percentages are by weight, all

temperatures are in degrees Centigrade, and pressure is at or near atmospheric pressure.

6.1 Sequence Alignment and Phylogenetic Analysis

PAP2 locus and its genomic organization, including its intron/exon boundaries, were identified in the Arabidopsis Col-0 ecotype (http://www.arabidopsis.org). Sequence alignment and phylogenetic tree were conducted using MEGA4 (Kumar *et al.*, 2004) and *ClustalW* program (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Amino acid sequence comparisons were performed using CLC Sequence Viewer 5.1.1 (www.clcbio.com).

Twenty nine PAP-like sequences were identified from the Arabidopsis genome and a phylogenetic tree was produced by neighbor-joining algorithm (FIG. 1). The gene locus of *AtPAP2* (At1g13900) composes of two exons and the coding region is 1971bp in length (SEQ ID NO: 1), which is predicted to encode a polypeptide of ~73.7-KD. Among the twenty nine PAP-like protein sequences, only AtPAP2 and AtPAP9 carry a unique hydrophobic motif at their C- termini by TMHMM analysis (http://www.cbs.dtu.dk/services/TMHMM-2.0/) (FIG. 3). AtPAP2 was found to share 72% sequence identity in amino acid sequence with AtPAP9. Two sequences from *Zea mays* (Accession No: ACG47621) and *Oryza sativa* (Accession No: BAC15853.1) were found to share 58% and 57% a.a. identity with AtPAP2, respectively. Their sequences were aligned in FIG. 2.

AtPAP2-like sequences from other plant species that carry a hydrophobic motif at their C- termini were retrieved by tblastn program from Plant GDB database (http://www.plantgdb.org/) and NCBI database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) using the amino acid sequence of AtPAP2 as the search sequence. cDNA and protein sequences that share high homology with that of AtPAP2 were identified in *Zea mays* (SEQ ID NOs: 7 and 8), *Brassica rapa* (SEQ ID NOs: 18 and 19), *Hordeum vulgare* (SEQ ID NOs: 20 and 21), *Medicago truncatula* (SEQ ID NOs: 22 and 23), *Physcomitrella patens* (SEQ ID NOs: 24 and 25), *Populus trichocarpa* (SEQ ID NOs: 26 and 27), *Saccharum officinarum* (SEQ ID NOs: 28 and 29), *Solanum tuberosum*

(SEQ ID NOs: 30 and 31), *Vitis vinifera* (SEQ ID NOs: 32 and 33), *Oryza sativa* (SEQ ID NOs: 34 and 35), *Gossypium hirsutum* (SEQ ID NOs: 36 and 37) *Panicum virgatum* (SEQ ID NOs: 38 and 39), *Solanum lycopersicum* (SEQ ID NOs: 40 and 41), *Sorghum bicolor* (SEQ ID NOs: 42 and 43) and *Triticum aestivum* (SEQ ID NOs: 44 and 45).

The cDNA sequences of AtPAP-like sequences were amplified from a local *Glycine max* variety (SEQ ID NO: 5) and the *Brassica napus* cultivar Westar (SEQ ID NO: 46) by RT-PCR using primers designed from corresponding EST sequences, which were retrieved from the Plant GDB database (http://www.plantgdb.org/).

6.2 Screening of T-DNA line and production of Overexpression lines and Complementation lines in *Arabidopsis*

T-DNA insertion lines of PAP2 gene (Arabidopsis genomic locus name: Salk_013567), in the Col ecotype were obtained from Arabidopsis Biological Resources Center (Alonso *et al.*, 2003). Homologous T-DNA lines were identified by genomic PCR screening from SIGnAl database (http://signal.salk.edu/cgi-bin/tdnaexpress) by using the primers (LBa1, 5'-TGGTTCACGTAGTGGGCCATCG-3', SEQ ID NO: 9) and PAP2 specific forward primer (P2LP, 5'- TTGAAGTTTAACATGCCTGGG-3, SEQ ID NO: 10) and reverse primer (P2RP, 5'-TCCAATGCTCGA TTGATTAGC-3', SEQ ID NO: 11). The PCR product was sequenced and the T-DNA insertion site was confirmed. To exclude the possibility that another T-DNA locus interferes with the PAP2 mutant site, homologous *pap2* mutant lines were backcrossed to the wild-type to dilute the potential T-DNA sites. The produced heterozygous *pap2* mutants were grown on the MS plates containing 50mg/ml Kanamycin. The ratio of the resistant to sensitive plants was about 3:1. These results demonstrated a single insertion locus site of the T-DNA line (*pap2-8*) lines.

The inability of the T-DNA line to express full length AtPAP2 mRNA was confirmed by RT-PCR. Total RNA was extracted from 10-day-old seedlings grown on MS with 2% (w/v) sucrose using the TRIzol RNA isolation method (Invitrogen) with DNase I treatment. cDNAs were generated using Superscript III reverse transcriptase (Invitrogen, Carlsbad, CA, USA) using an oligo dT primer. Two gene-specific primers, P2YF (5'-GGCCGTCGACATGATCGTTAATT TCTCTTTC -3' SEQ ID NO: 12) and P2NR (5'- CCGGACTAGTTCATGTCTCCTCGTTCTTGAC -

3' SEQ ID NO: 13), were used to amplify a 1971bp coding region of AtPAP2. For each sample, 1 μg of cDNA was amplified for 30 cycles, with an annealing temperature of 50°C and using elongation factor (EF) primers, EF-1 (5'-GTTTCACATCAACATTGTGGTCA TTGG -3, SEQ ID NO: 14) and EF- 2 (5'-GAGTACTTGGGGGTAGTGGCATCC-3, SEQ ID NO: 15) (Axelos *et al.*, 1989) for control experiment.

The inability of the T-DNA line to express protein was confirmed by Western blotting analysis (FIG. 4). Antiserum specific to AtPAP2 was raised in rabbit as described in Section 6.3.

To create transgenic AtPAP2 overexpressing lines or expressing this gene in the knockout mutants, the full length coding region of the AtPAP2 cDNA was amplified by PCR using primers P2YF (SEQ ID NO: 12) and P2NR (SEQ ID NO: 13). A *Sal*I site and a *Spe*I site were engineered into P2YF and P2NR, respectively. The resulting product (1976bp) was inserted into the XhoI/Spe I sites of a binary vector, immediately downstream to the cauliflower mosaic virus (CaMV) 35S promoter (FIG. 5).

The vector was introduced into *Agrobacterium tumefaciens* strain GV3101 and then transformed by the floral dip method (Clough and Bent, 1998), into wild-type Col-0 to generate PAP2-overexpressing lines or into homologous *pap2* plants (T-DNA lines) to generate complementation lines. Through 2 generations of selection on MS agar plate with 50 mg/l Basta (Riedel-deHaen), homologous 35S:PAP2 transgenic lines were obtained. The resistant plants were transferred to soil to grow to maturity, and their transgenic status was further confirmed by PCR and immunoblot analyses. As shown in FIG. 6A, AtPAP2 protein was overexpressed in OE lines but was absence from the T-DNA line. The homozygous T3 seeds of the transgenic plants were used for further analysis.

To create transgenic AtPAP15 overexpression lines, the cDNA of AtPAP15 was also amplified by RT-PCR and then subcloned into a plant binary vector which bared a kanamycin-resistant gene and a cauliflower mosaic virus 35S promoter (CaMV). This expression construct named was then mobilized into *Agrobacterium tumefaciens* strain EHA105 by freeze-thaw transformation (Hofgen and Willmitzer, 1988) and transformed into Arabidopsis. Transgenic status was further confirmed by PCR and immunoblot analyses using an anti-AtPAP15 antiserum. As shown in FIG.

6B, AtPAP15 protein was overexpressed in OE lines. The homozygous T3 seeds of the transgenic plants were used for further analysis.

6.3 Production of PAP2 Polycolonal Antiserum and Western Blots Analysis

A fragment of *AtPAP2* cDNA corresponding to the N terminal 120 amino acids (from 21 to 141) was amplified using forward primer P2AF (5'-GGTTGAGCTCGATTCTAAAGCGACCATTTC-3', SEQ ID NO: 16) and reverse primer P2AR (5'- TTTTGGTACC**TCA**GGATCCGAA AGTCAGC-3', SEQ ID NO: 17). The PCR product was cleaved by *Sac*I and *Kpn*I and cloned into the pRsetA vector (Invitrogen) so that the coding sequence of the first 120 a.a. of AtPAP2 was fused to a His-tag sequence. The resulting plasmid was transformed into *Escherichia coli* strain BL21 (DE3). The BL21 cells were induced at 30°C by 0.1 mM isopropylthio-ß-D- galactoside for 4 h and resuspended in 100 mM NaCl and 50mM Tris-HCl, pH 7.5, 2mM phenylmethylsulphonyl fluoride (PMSF). The lysates were sonicated 5 times for 30 s each. The overexpressed His-AtPAP2 fusion proteins in inclusion bodies were centrifuged at 5000 x *g* for 15 min, and the pellets were solubilized in 150 mM NaCl, 8 M urea, and 20 mM Tris-HCl, pH 7.5. The fusion proteins were purified on a HisTrap FF (GE Healthcare) column and were used for standard immunization protocols in rabbits.

6.4 Expression analysis of AtPAP2 mRNA and its protein levels

The mRNA expression level of AtPAP2 was analyzed by the Spot History program (http://affymetrix.arabidopsis.info/narrays/spothistory.pl) that presented the expression levels of a given gene in thousands of microarray (Affymetrix ATH1 microarray) database. Spot history analysis indicated that the expression of AtPAP2 was constitutive but is relative low in most experimental circumstances (FIG. 7a). To determine AtPAP2 expression levels, different tissues of wild-type *A. thaliana* (Col-0) were collected.

The expression level of proteins were also studied by western blotting, using the anti-AtPAP2 antiserum generated from Section 6.3. Total plant soluble protein was extracted from wild-type *A.thaliana*, T-DNA line, AtPAP2-overexpress lines in grinding buffer (Tris-HCl 50mM, pH7.4 containing 150mM NaCl, 1mM EDTA, 0.2 mM PMSF) on ice. Protein extracts were centrifuged at 16000 x g and supernatants were collected for Bradford protein concentration determination assay. Equal amount

of protein samples (50-90 μg/lane in different experiments) were loaded and separated in 12% (w/v) SDS-PAGE. The separated proteins were transferred to Hybond C-Extra membranes (Amersham Biosciences) (400mA, 1h). Membranes were blocked with 5% (w/v) non-fat milk in TTBS washing buffer (pH 7.6) for 2 hours and probed with specific anti-AtPAP2 antiserum for 3 hours or overnight at an 1: 1000 dilution at 4℃. After rinsing the membrane with three changes of TTBS washing buffer (20mM Tris-HCl, pH7.6, 136mM NaCl, 0.1% Tween20) in half an hour, HRP-labeled secondary antibody, diluted 1:10,000 in TTBS washing buffer was added. After 2 hours, the membrane was washed thrice before the bands were visualized by Enhanced Chemiluminescence method (Amersham Biosciences). As shown in FIG. 7b, AtPAP2 protein was expressed in all tissues tested (Leaf, Flower, Stem, Root, Silique) at equal levels. The protein expression level of AtPAP2 during germination was very stable too (FIG. 7c) and was independent of phosphorus status (FIG. 7d).

6.5 Growth Phenotypes of WT, T-DNA line and OE lines

Arabidopsis seeds were soaked in water at 4°C for 3 days. The seeds were surface sterilized and sown on Murashige and Skoog (MS) medium supplemented with 2% (w/v) sucrose for 10 days. Seedlings with 2 rosette leaves of the same size were transferred to soil under Long Day (16h light at 22°C/ 8h dark at18°C) or Short Day (8h light at 22°C/ 16h dark at18°C) conditions in a plant growth chamber. Flowering time was started to be measured by scoring the number of rosette leaves and cauline leaves when the primary inflorescence florescence reached 1 cm above the rosette leaves. Ten to 20 plants were scored for each line (Liu *et al.*, 2008; Wu *et al.*, 2008).

The inflorescences of OE lines of AtPAP2 emerged earlier (5-6 days for Long Day, 14-16 days for Short Day) than that of the WT and T-DNA lines (Table 1). Under Long Day conditions, the number of rosette leaves of the OE lines were less (5-6 leaves) than the WT during the emergence of inflorescence (Table 1 and Fig. 8). At day 28 (Long Day), the OE lines of AtPAP2 had more cauline leaves and inflorescences than the WT and T-DNA lines, but had less rosette leaves (Table 2.). This phenotype observation was repeated at least four times and the results of one of the experiments were shown here.

Table 1. AtPAP2 OE lines flowered at an earlier developmental stage.

	Long Day (16h/8h)				Short Day (8h/16h)			
Lines	AEI	SD	NRL	SD	AEI	SD	NRL	SD
Col-0	26.9	1.2	13.0	0.8	41.0	4.7	18.0	3.0
T-DNA	25.7	0.7	11.6	1.1	40.7	4.9	15.0	3.0
OE7	20.0*	1.1	6.4*	0.5	25.6*	1.3	5.3*	0.5
OE21	20.8*	0.6	6.5*	0.7	26.0*	1.1	5.4*	0.5

AEI: Average date of emergence of inflorescence
NRL: No. of rosette leaves at the first appearance of inflorescence
* Statistically ($p<0.001$) different from the wild-type (n=15).

Table 2. Phenotypes of AtPAP2 OE lines at Day 28 (Long Day).

Lines	No. of Rosette Leaf	SD	No. of Cauline Leaf	SD	No. of Inflorescence	SD
Col-0	14.5	1.2	1.6	0.5	1.0	0.0
T-DNA	16.7	1.7	1.9	0.6	1.0	0.0
OE7	9.9*	1.0	6.0*	1.2	3.6*	0.7
OE21	10.2*	1.8	7.2*	1.6	3.7*	1.1

* Statistically ($p<0.001$) different from the wild-type (n=15).

At maturity (Long Day), the number of siliques and the total weight of seeds harvested from each line were recorded. Two separate experimental trials are shown in Tables 3A and 3B. Our results showed that overexpression of AtPAP2 resulted in increase number of siliques per plant and the seed yield per plant. Compared to that of the wild-type, the seed yield of the two overexpression lines shown in Table 3A increased 38-40%. Compared to that of the wild-type, the seed yield of the two overexpression lines shown in Table 3B increased 54-58%.

Table 3A. OE lines produced more siliques and seeds (Trial 1).

Lines	No. of siliques/plant	SD	Weight of seeds (g)/plant	SD	N
Col-0	327.4	53.3	0.188	0.047	5
T-DNA	236.6*	60.2	0.121*	0.040	7
OE7	453.2**	62.1	0.264**	0.039	5
OE21	498.2#	52.5	0.260**	0.049	7

Statistically ($p<0.02$*, $p<0.01$**, $p<0.001$#) different from the wild-type.

Table 3B. OE lines produced more siliques and seeds (Trial 2).

Lines	No. of siliques/plant	SD	Weight of seeds (g)/plant	SD	N
Col-0	396.4	89.5	0.225	0.058	13
T-DNA	386.3	70.4	0.240	0.049	12
OE7	610.9*	76.6	0.351*	0.050	7
OE21	624.9*	94.7	0.355*	0.066	11

Statistically ($p<0.0001$*) different from the wild-type.

However, the OE lines of AtPAP15 grew normally and were not different from the wild-types. Therefore, the enhanced growth performance was due to the overexpression of AtPAP2, which bears a transmembrane-like motif at its C-terminus (FIGS. 2 and 3).

6.6 Growth Phenotypes of Truncated AtPAP2 Constructs

An alternate vector construct employing the sequence for AtPAP2 was also constructed using analogous techniques to those described above. As shown in Figure 12, a construct equivalent to the OE lines of AtPAP2 missing the C-terminal motif (residues 614-636 of SEQ ID NO: 2) was constructed(P2C lines). Transgenic plants were generated using substantially identical techniques to those described above. Western blot analysis was used to confirm the over-expression of the AtPAP2 fragment proteins in transformed plant lines. Performance of Western blot analysis was identical to that reported above. As show in Figure 13, the P2C lines were strongly overexpressed. The growth phenotype of the P2C lines appeared to be indifferent from the wild-type, which is indicative of the importance of the C-terminal domain of AtPAP2 in developing an increased growth phenotype.

6.7 MS/MS analysis of sucrose and glucose levels in leaf

Rosette leaves of plants of various developmental stages were harvested at the end of the light period of 21-day-old plants. Soluble sugars were extracted from Arabidopsis using chloroform/methanol method (Lunn *et al.*, 2006; Antonio *et al.*, 2007; Luo *et al.*, 2007). 100 mg plant tissues were ground to a fine powder in liquid nitrogen and mixed and vortexed with 250µl ice-cold chloroform: methanol (3:7, v/v). Soluble metabolites were then extracted at -20 °C overnight. 200µl water was added to the mixture with repeated shaking. The extracts were centrifuged at 16000 x g for 10min and the supernatant was collected. The pellet was re-extracted by 200µl water and the supernatant was collected by centrifugation as described above. The combined supernatant was evaporated to dryness using a SpeedVac and the pellet was re-dissolved in 200µl water. Finally, debris was removed by centrifugation at 16000 x g for 30min.

20µl filtered samples were analyzed by an API-3000 triple-quadrupole mass spectrometer (Applied Biosystems) via an electrospray ionization source. The

parameters, optimized by 0-40μg/ml glucose and sucrose standards, were as following: curtain gas (CUR) 25, nebulizer gas (GS1) 50, auxiliary gas (GS2) 30, ionspray voltage -4.5k V, temperature 400 °C, declustering potential (DP) -106 V, entrance potential (EP) -8.5 V, collision cell entrance potential (CEP) -46.7 V, collision energy (CE) 20 V. The peaks were identified by comparison with glucose and sucrose standards and the amount of sugars were quantified by standard curves of these sugars. The Analyst 1.3.1 software (Applied Biosystems) was used for data acquisition, peak integration, and calculation. The amount of sucrose and glucose at the end of day in the shoots of 21-day-old soil grown plants were shown in FIG. 9. It was found that the levels of both sugars were significantly higher than that in WT.

6.8 Recovery of plants after prolonged darkness treatment

Seeds of wild-type, T-DNA, OE7 and OE21 lines were germinated in MS (2% sucrose) medium for 10 days. Seedlings with 2 small visible rosette leaves (1mm) of the same size were transferred to soil for another 12 days in normal growth conditions (LDs, 16h/light (22°C)/8h darkness (18°C)). The light source of the growth chamber was then switched off for 12 days. Then the plants were allowed to recover under the 16h/light (22°C)/8h darkness (18°C) cycle for 10 days. The plants that stayed green and that continued to emerge inflorescence were recorded in Table 4.

Table 4. Surviving rate and flowering ratio after prolonged darkness treatment.

	Flowering after recovery	Recovery (leaf greening)
Wild-type	8/12	11/12
T-DNA	5/12	8/12
OE7	9/9	9/9
OE21	12/12	12/12

Extended darkness could induce carbohydrate starvation (Thompson *et al.*, 2005). Our data showed that the OE lines exhibited 100% recovery rate under prolonged (12 days) darkness treatment, which was higher than that of the WT and the T-DNA line (FIG. 10). This could be attributed to a higher endogenous sugar levels (FIG. 9) in the OE lines.

6.9 Phenotypes of plants under NaCl and ABA treatments

5-day-old seedlings grown on MS agar were transferred to MS agar with NaCl (50mM, 100mM, 150mM), ABA (0.1uM, 0.2uM. 0.5uM, 1uM, 2uM) or sorbitol (300mM, 400mM, 500mM). Alternatively, seeds were directly germinated on the treatment media. Wild-type, T-DNA and OE lines did not show remarkable phenotypic differences under the above conditions.

6.10 Subcellular Fractionation

Rosette leaves of three-week-old wild-type (Col-0) *Arabidopsis* were harvested and stored at –80 °C freezer until use. Tissues (4-5 g) were ground to fine powder in liquid nitrogen using a mortar with a pestle. The powder was transferred into 10 ml grinding buffer (0.3 M sucrose, 40 mM Tris-HCl (pH 7.8), 5 mM $MgCl_2$, 1mM PMSF) and swelled on ice for 5 min. Homogenization was performed for two 30-second pulses at high-speed setting. The homogenate was filtered through two layers of Miracloth (Tetko, Elmsford, N.Y., USA). Subsequently, the homogenate was separated by centrifugation at 350 *g* for 10 min at 4 °C. The pellet (crude nuclear) was further layered onto 1 ml of 2.3 M sucrose, 50 mM Tris-HCl (pH 8.8), 5 mM $MgCl_2$ in an Eppendorf tube for centrifugation at 15,000 *g* 10 min at 4 °C, to obtain the nuclear fraction in the derived pellet. Supernatants from the first low-speed centrifugation (350 *g*) were centrifuged at 12,000 x *g* for 20 min at 4 °C. The pellet contained large particles including mitochondria, chloroplasts and peroxisomes. The supernatant was further centrifuged at 100,000 x g for 1 h at 4 °C to yield the soluble cytosol fraction in the resulting supernatant. The pellet representing the membrane fraction was resuspended in 0.1 ml grinding buffer. Protein concentration in the extract was determined following the method of Bradford (Bradford, 1976) using the Bio-Rad Protein Assay Kit I.

To isolate cell wall, leaf tissues were homogenized in grinding buffer (62.5 mM Tris-HCl, pH 7.5, 5 mM DTT, 1% (v/v) bovine serum albumin, 2 mM phenylmethylsulfonyl fluoride, 2 μg/ml leupeptin, 2 μg/ml E-64, 2 μg/ml pepstatin A) using a Polytron (full speed, 3 × 10 s). The homogenate was centrifuged at 1,000 × *g* for 3 min. The pellet was washed with ice-cold grinding buffer (without 1% BSA) 10 times. Finally the (cell wall) pellet was washed by resuspending in 500 mM $CaCl_2$,

20 mM NaCl, 62.5 mM Tris-HCl, pH 7.5, and spinning at 10,000 × *g* for 15 min (He *et al.*, 1996).

The subcellular fractions were run in a SDS-PAGE gel and were probed with anti-AtPAP2 antiserum. AtPAP2 was detected in membrane and soluble protein fractions but not in nucleus, mitochondria nor chloroplasts (FIG. 11).

In summary, Arabidopsis plants transformed with the AtPAP2 gene have the following phenotypes when they were compared with the wild-type: (1) Faster growth rate (Tables 1 and 2); (2) Higher sucrose content (FIG. 9); (3) Higher glucose content (FIG. 9); and (4) Higher crop yield (Table 3).

Those skilled in the art will recognize, or be able to ascertain many equivalents to the specific embodiments described herein using no more than routine experimentation. Such equivalents are intended to be encompassed by the following claims.

With respect to any figure or numerical range for a given characteristic, a figure or a parameter from one range may be combined with another figure or a parameter from a different range for the same characteristic to generate a numerical range.

All publications, patents and patent applications mentioned in this specification are herein incorporated by reference into the specification to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated herein by reference.

Citation or discussion of a reference herein shall not be construed as an admission that such is prior art to the present patent application.

While the embodiments have been explained in relation to certain embodiments, it is to be understood that various modifications thereof will become apparent to those skilled in the art upon reading the specification. Therefore, it is to be understood that the embodiments disclosed herein are intended to cover such modifications as fall within the scope of the appended claims. Features of two or more of any of the above embodiments can be combined to form additional embodiments.

Other than in the operating examples, or where otherwise indicated, all numbers, values and/or expressions referring to quantities of ingredients, reaction conditions, etc., used in the specification and claims are to be understood as modified in all instances by the term "about."

6.11 Assays of enzymes involved in sucrose metabolism

Sucrose phosphate synthesis (SPS), sucrose synthesis (SuSy), cytosolic invertase and cell wall invertase activities in the shoot of 20-day-old plants were determined. Samples were collected 8h after the light and dark period (Long Day). SPS activity was measured under both optimal (Vmax) and limiting (V limit) assay conditions (Park *et al.*, 2008). SuSy, cytosolic invertase and insoluble cell wall invertase activities were also determined (Doehlert, 1987). The assays were repeated three times and the SPS (Vmax and V limit) activities of both independent lines were significantly higher than that of the wild-type and T-DNA lines in all three repeated experiments. The data of a representative experiment is shown in table 5. In contrast to SPS, SuSy, cytosolic invertase and cell wall invertase activities were not different among the lines.

Table 5. Enzyme assays

Plant line	WT	T-DNA	OE7	OE21
Sucrose phosphate synthase (μM sucrose/μg enzyme extracts/hour)				
Vmax (Day)	108.8 ± 18.1	116.1 ± 11.9	***157.9 ± 21.7****	***159.5 ± 19.0****
Vlimit (Day)	63.6 ± 6.4	69.5 ± 2.5	***90.8 ± 18.5****	***79.8 ±13.7****
Vmax (Night)	118.3 ± 11.4	104.9 ± 14.4	***150.9 ± 19.4****	***136.5 ± 15.1****
Vlimit (Night)	74.9 ±6.3	56.7 ± 3.3	***93.4 ± 3.6****	***97.3 ± 10.9****
Sucrose synthase (μM glucose /μg enzyme extracts/hour)				
Day	249.2 ± 4.6	247.0 ± 24.0	248.6 ± 4.5	255.2 ± 5.4
Night	249.4 ± 7.2	252.7 ± 8.8	250.8 ± 8.7	258.8 ± 5.5
Cytosolic invertase (μM glucose /μg enzyme extracts/hour)				
Day (Acid)	14.3 ± 2.3	12.1 ± 5.4	18.3 ± 9.1	18.0 ± 0.8
Night (Acid)	14.0 ± 6.8	26.1 ± 10.9	17.6 ± 4.3	13.8 ± 0.8
Day (Alkaline)	169.7 ± 9.8	161.2 ± 32.3	160.9 ± 27.9	178.8 ± 16.9
Night (Alkaline)	130.0 ± 8.2	105.1 ± 12.6	136.1 ± 13.6	136.6 ± 1.4
Cell wall invertase (μM sucrose/μg enzyme extracts/hour)				
Day (Acid)	22.3 ± 3.9	18.0 ± 4.9	23.5 ± 6.3	24.0 ± 6.6
Night (Acid)	21.6 ± 10.1	17.0 ± 4.3	28.3 ± 4.2	22.3 ± 3.3
Day (Alkaline)	121.7 ± 2.8	127.1 ± 2.2	101.8 ± 6.4	105.5 ± 2.0
Night (Alkaline)	138.9 ± 6.3	151.5 ± 8.7	123.4 ± 5.6	135.4 ± 14.7

(** $P<0.01$; * $P<0.05$)

ARTICLES

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. and Ecker, J.R. (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science*, **301**, 653-657.

Antonio, C., Larson, T., Gilday, A., Graham, I., Bergstrom, E. and Thomas-Oates, J. (2007) Quantification of sugars and sugar phosphates in Arabidopsis thaliana tissues using porous graphitic carbon liquid chromatography-electrospray ionization mass spectrometry. *J Chromatogr A*, **1172**, 170-178.

Axelos, M., Bardet, C., Liboz, T., Le Van Thai, A., Curie, C. and Lescure, B. (1989) The gene family encoding the Arabidopsis thaliana translation elongation factor EF-1 alpha: molecular cloning, characterization and expression. *Mol Gen Genet*, **219**, 106-112.

Bozzo, G.G., Raghothama, K.G. and Plaxton, W.C. (2002) Purification and characterization of two secreted purple acid phosphatase isozymes from phosphate-starved tomato (Lycopersicon esculentum) cell cultures. *Eur.J Biochem*, **269**, 6278-6286.

Bozzo, G.G., Raghothama, K.G. and Plaxton, W.C. (2004) Structural and kinetic properties of a novel purple acid phosphatase from phosphate-starved tomato (Lycopersicon esculentum) cell cultures. *Biochem J*, **377**, 419-428.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, **72**, 248-254.

Cashikar, A.G., Kumaresan, R. and Rao, N.M. (1997) Biochemical Characterization and Subcellular Localization of the Red Kidney Bean Purple Acid Phosphatase. *Plant Physiol*, **114**, 907-915.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J*, **16**, 735-743.

Coello, P. (2002/11) Purification and characterization of secreted acid phosphatase in phosphorus-deficient Arabidopsis thaliana. *Physiologia Plantarum*, **116**, 293-298.

del Pozo, J.C., Allona, I., Rubio, V., Leyva, A., de la, P.A., Aragoncillo, C. and Paz-Ares, J. (1999) A type 5 acid phosphatase gene from Arabidopsis thaliana is induced by phosphate starvation and by some other types of phosphate mobilising/oxidative stress conditions. *Plant J.*, **19**, 579-589.

Doehlert, D.C. (1987) Ketose Reductase Activity in Developing Maize Endosperm. *Plant Physiol*, **84**, 830-834.

He, Z.H., Fujiki, M. and Kohorn, B.D. (1996) A cell wall-associated, receptor-like protein kinase. *J Biol Chem*, **271**, 19789-19793.

Hegeman, C.E. and Grabau, E.A. (2001) A novel phytase with sequence similarity to purple acid phosphatases is expressed in cotyledons of germinating soybean seedlings. *Plant Physiol.*, **126**, 1598-1608.

Hofgen, R. and Willmitzer, L. (1988) Storage of competent cells for Agrobacterium transformation. *Nucleic Acids Res*, **16**, 9877.

Kaida (2003) Isolation and characterization of four cell wall purple acid phosphatase genes from tobacco cells. *Biochim.biophys acta*, **1625**, 134-140.

Kim, S. and Gynheung, A. (1996) Isolation and characterization of a pollen-specific cDNA clone from Easter lily. *J. Plant Biol.* , **39**, 197-202.

Klabunde, T., Strater, N., Frohlich, R., Witzel, H. and Krebs, B. (1996) Mechanism of Fe(III)-Zn(II) purple acid phosphatase based on crystal structures. *J.mol.biol.*, **259**, 737-748.

Klabunde, T. and Krebs, B. (1997) The dimetal center in purple acid phosphatases. *Metal Sites in Proteins and Models*, **89**, 177-198.

Kumar, S., Tamura, K. and Nei, M. (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform*, **5**, 150-163.

Li, D., Zhu, H., Liu, K., Liu, X., Leggewie, G., Udvardi, M. and Wang, D. (2002) Purple acid phosphatases of Arabidopsis thaliana. Comparative analysis and differential regulation by phosphate deprivation. *J.Biol.Chem.*, **277**, 27772-27781.

Li, D. and Wang, D. (2003) cDNA cloning and in vitro expression of three putatave purple acid phosphatase genes from Arabidopsis. *Journal of nature science Hunan Normal University*, **26**, 78-82.

Liao, H., Wong, F.L., Phang, T.H., Cheung, M.Y., Li, W.Y., Shao, G., Yan, X. and Lam, H.M. (2003) GmPAP3, a novel purple acid phosphatase-like gene in soybean induced by NaCl stress but not phosphorus deficiency. *Gene*, **318**, 103-111.

Liu, L.J., Zhang, Y.C., Li, Q.H., Sang, Y., Mao, J., Lian, H.L., Wang, L. and Yang, H.Q. (2008) COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. *Plant Cell*, **20**, 292-306.

Lung, S.C., Leung, A., Kuang, R., Wang, Y., Leung, P. and Lim, B.L. (2008) Phytase activity in tobacco (Nicotiana tabacum) root exudates is exhibited by a purple acid phosphatase. *Phytochemistry*, **69**, 365-373.

Lunn, J.E., Feil, R., Hendriks, J.H., Gibon, Y., Morcuende, R., Osuna, D., Scheible, W.R., Carillo, P., Hajirezaei, M.R. and Stitt, M. (2006) Sugar-induced increases in trehalose 6-phosphate are correlated with redox activation of ADPglucose pyrophosphorylase and higher rates of starch synthesis in Arabidopsis thaliana. *Biochem J*, **397**, 139-148.

Luo, B., Groenke, K., Takors, R., Wandrey, C. and Oldiges, M. (2007) Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography-mass spectrometry. *J Chromatogr A*, **1147**, 153-164.

Park, J.Y., Canam, T., Kang, K.Y., Ellis, D.D. and Mansfield, S.D. (2008) Over-expression of an arabidopsis family A sucrose phosphate synthase (SPS) gene alters plant growth and fibre development. *Transgenic Res*, **17**, 181-192.

Patel, K., Lockless, S., Thomas, B. and McKnight, T.D. (1998) Secreted purple acid phosphatase from Arabidopsis thaliana. *American Society of Plant Physiogists*, 373-374.

Schenk, G., Ge, Y., Carrington, L.E., Wynne, C.J., Searle, I.R., Carroll, B.J., Hamilton, S. and de-Jersey, J. (1999) Binuclear metal centers in plant purple acid phosphatases: Fe-Mn in sweet potato and Fe-Zn in soybean. *Arch.Biochem Biophys*, **370**, 183-189.

Schenk, G., Korsinczky, M.L., Hume, D.A., Hamilton, S. and DeJersey, J. (2000) Purple acid phosphatases from bacteria: similarities to mammalian and plant enzymes. *Gene*, **255**, 419-424.

Shimaoka, T., Ohnishi, M., Sazuka, T., Mitsuhashi, N., Hara-Nishimura I, Shimazaki, K., Maeshima, M., Yokota, A., Tomizawa, K. and Mimura, T. (2004) solation of intact vacuoles and proteomic analysis of tonoplast from suspension-cultured cells of Arabidopsis thaliana. *Plant Cell Physiol*, **45**, 672-683.

Thompson, A.R., Doelling, J.H., Suttangkakul, A. and Vierstra, R.D. (2005) Autophagic nutrient recycling in Arabidopsis directed by the ATG8 and ATG12 conjugation pathways. *Plant Physiol*, **138**, 2097-2110.

Veljanovski, V., Vanderbeld, B., Knowles, V.L., Snedden, W.A. and Plaxton, W.C. (2006) Biochemical and molecular characterization of AtPAP26, a vacuolar purple acid phosphatase up-regulated in phosphate-deprived Arabidopsis suspension cells and seedlings. *Plant Physiol*, **142**, 1282-1293.

Wu, J.F., Wang, Y. and Wu, S.H. (2008) Two New Clock Proteins, LWD1 and LWD2, Regulate Arabidopsis Photoperiodic Flowering. *Plant Physiol*, **148**, 948-959.

Wu, P., Ma, L., Hou, X., Wang, M., Wu, Y., Liu, F. and Deng, X.W. (2003) Phosphate starvation triggers distinct alterations of genome expression in Arabidopsis roots and leaves. *Plant Physiol*, **132**, 1260-1271.

Xiao, K., Harrison, M.J. and Wang, Z.Y. (2005) Transgenic expression of a novel M. truncatula phytase gene results in improved acquisition of organic phosphorus by Arabidopsis. *Planta*, **222**, 27-36.

Xiao, K., Katagi, H., Harrison, M. and Wang, Z.Y. (2006) Improved phosphorus acquisition and biomass production in Arabidopsis by transgenic expression of a purple acid phosphatase gene from M. truncatula. *Plant Science*, **170**, 191-202.

Zhang, W., Gruszewski, H.A., Chevone, B.I. and Nessler, C.L. (2008) An Arabidopsis purple Acid phosphatase with phytase activity increases foliar ascorbate. *Plant Physiol*, **146**, 431-440.

Zhu, H.F., Qian, W.Q., Lu, X.Z., Li, D.P., Liu, X., Liu, K.F. and Wang, D.W. (2005) Expression patterns of purple acid phosphatase genes in Arabidopsis organs and functional analysis of AtPAP23 predominantly transcribed in flower. *Plant Molecular Biology*, **59**, 581-594.

CLAIMS

What is claimed is:

1. A method to make a transgenic plant having increased rate of plant growth and elevate plant yields comprising:

introducing a gene coding for a phosphatase into a plant, the phosphatase gene encodes for a polypeptide comprising a C-terminal motif having the sequence of SEQ ID NO: 66 or a homologue thereof, and one or more sequences selected from the group consisting of SEQ ID NOS: 48-53.

2. The method of claim 1, wherein the phosphatase gene is a purple acid phosphatase gene comprising a nucleotide sequence encoding a polypeptide comprising an amino acid sequence of one or more selected from SEQ ID NOS: 2, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35,and 47, and homologues thereof.

3. The method of claim 1, wherein the phosphatase gene is a purple acid phosphatase gene comprising a nucleotide sequence selected from SEQ ID NOS: 1, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, and 46 and homologues thereof.

4. The method of claim 2, wherein said nucleotide sequence comprises one or more selected from the group consisting of SEQ ID NOs 1, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, and 46 and homologues thereof.

5. The method of claim 2, wherein said nucleotide sequence comprises a nucleotide sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 2 and conservatively substituted variants thereof.

6. The method of claim 2,wherein said nucleotide sequence comprises a nucleotide sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 54 and 55 and conservatively substituted variants thereof.

7. The method of claim 6, wherein said nucleotide sequence comprises a nucleotide sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 50 and 51.

8. The method of claim 2, wherein said nucleotide sequence comprises a nucleotide sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 57-60 and conservatively substituted variants thereof.

9. The method of claim 8, wherein said nucleotide sequence comprises a nucleotide sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 50 and 51.

10. The method of claim 2, wherein said nucleotide sequence comprises a nucleotide sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 61-64.

11. The method of claim 1, wherein introducing the phosphatase gene up-regulates the enzymatic activity of sucrose phosphate synthase in the transgenic plant relative to the wild-type plant.

12. The method of claim 1, wherein introducing the phosphatase gene up-regulates the sucrose and/or glucose level in the transgenic plant relative to the wild-type plant.

13. The method of claim 1, wherein introducing the phosphatase gene increases the growth rate of the transgenic plant relative to the wild-type plants.

14. The method of claim 1, wherein introducing the phosphatase gene results in a higher crop yield of the transgenic plant relative to the wild-type plants.

15. The method of claim 1, wherein the plant is a species selected from one of the group consisting of the following genera: Asparagus, Bromus, Hemerocalli, Hordeum, Loliu, Panicum, Pennisetum, Saccharum, Sorghum, Trigonell, Triticum, Zea,

Antirrhinum, Arabidopsis, Arachis, Atropa, Brassica, Browallia, Capsicum, Carthamus, Cichorium, Citrus, Chrysanthemum, Cucumis, Datura, Daucus, Digitalis, Fragaria, Geranium, Glycine, Helianthus, Hyscyamus, Ipomoea, Latuca, Linum, Lotus, Solanum lycopersicon, Majorana, Malva, Gossypium, Manihot, Medicago, Nemesia, Nicotiana, Onobrychis, Pelargonium, Petunia, Ranunculus, Raphanus, Salpiglossis, Senecio, Sinapis, Solanum, Trifolium, Vigna, and Vitis.

16. The method of claim 1, wherein the plant is a species selected from the family Brassica.

17. A transformed plant, comprising:

a plant comprising at least one additional gene coding for a phosphatase relative to a corresponding wild-type plant, the phosphatase gene encodes for a polypeptide comprising a C-terminal motif having the sequence of SEQ ID NO: 66 or a homologue thereof, and one or more sequences selected from the group consisting of SEQ ID NOS: 48-53.

18. The transformed plant of claim 17, wherein the phosphatase gene is a purple acid phosphatase gene comprising a nucleotide sequence encoding a polypeptide comprising an amino acid sequence of one or more selected from SEQ ID NOS: 2, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35,and 47, and homologues thereof.

19. The transgenic plant of claim 17, wherein the phosphatase gene is a purple acid phosphatase gene comprising a nucleotide sequence selected from SEQ ID NOS: 1, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, and 46 and homologues thereof.

20. The transgenic plant of claim 17, wherein the phosphatase gene is a purple acid phosphatase gene comprising a nucleotide sequence selected from SEQ ID NO: 1 or a homologue thereof.

21. The transformed plant of claim 17, wherein the phosphatase gene is a purple acid phosphatase gene comprising a nucleotide sequence encoding a polypeptide comprising an amino acid sequence of one or more selected from SEQ ID NOS: 2 or a homologue thereof.

22. The transformed plant of claim 17, wherein said plant is of a monocotyledonous species.

23. The transformed plant of claim 17, wherein said plant is of a dicotyledonous species.

24. The transformed plant of claim 17, wherein the plant is a species selected from one of the group consisting of the following genera: Asparagus, Bromus, Hemerocalli, Hordeum, Loliu, Panicum, Pennisetum, Saccharum, Sorghum, Trigonell, Triticum, Zea, Antirrhinum, Arabidopsis, Arachis, Atropa, Brassica, Browallia, Capsicum, Carthamus, Cichorium, Citrus, Chrysanthemum, Cucumis, Datura, Daucus, Digitalis, Fragaria, Geranium, Glycine, Helianthus, Hyscyamus, Ipomoea, Latuca, Linum, Lotus, Solanum lycopersicon, Majorana, Malva, Gossypium, Manihot, Medicago, Nemesia, Nicotiana, Onobrychis, Pelargonium, Petunia, Ranunculus, Raphanus, Salpiglossis, Senecio, Sinapis, Solanum, Trifolium, Vigna, and Vitis.

25. The transformed plant of claim 17, wherein the plant is a species selected from the family Brassica.

26. A vector comprising a plasmid comprising a phosphatase gene comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, and 46 and homologues thereof, wherein the plasmid is capable of transforming a bacterial cell.

27. The vector of claim 26, wherein the bacterial cell is Agrobacterium tumefaciens.

28. The vector of claim 26, wherein said phosphatase gene encodes an amino acid sequence comprising one or more selected from the group consisting of SEQ ID NOS: 2, 6, 8, 19, 21, 23, 25, 27, 29, 31, 33, 35, and 47 and homologues thereof, except that all or an N-terminal portion of amino acid residues 1 to 50 of these amino acid sequences are replaced by a plant signal peptide such that said polypeptide is sorted to

various organelles or compartments of plant cells upon expression of the phosphatase gene in a transformed host plant cell.

29. The vector of claim 26, wherein the phosphatase gene comprises a nucleic acid sequence selected from SEQ ID NO: 1 or a homologue thereof.

30. A host cell comprising the vector of claim 26.

31. The host cell of claim 30, wherein the phosphatase gene comprising the vector is operably linked to a heterologous promoter.

32. The host cell of claim 31, wherein the heterologous promoter is a plant promoter.

33. The host cell of claim 31, wherein the heterologous promoter is a promoter derived from cauliflower mosaic virus.

34. The host cell of claim 30, wherein the host cell is a plant species.

35. A method for preparing a cell or progeny thereof capable of expressing a purple acid phosphatase in a host cell, comprising:

transforming a bacterial cell with the vector of claim 26 and

transforming the host cell by transfer of DNA from the bacterial cell to the host cell.

36. The method of claim 35, wherein the bacterial cell is Agrobacterium tumefaciens and the host cell is a plant cell.

37. The method of claim 36, wherein the host cell is from the family Brassica.

38. An expression cassette comprising a phosphatase gene comprising a nucleotide sequence encoding an enzyme having phosphatase activity, wherein said nucleotide sequence hybridizes to a nucleotide sequences selected from the group consisting of SEQ ID NOS: 1, 5, 7, 18, 20, 22, 24, 26, 28, 30, 32, 34, and 46, under

stringent condition and is operably linked to regulatory nucleotide sequences such that said regulatory nucleotide sequences cause expression of the nucleotide sequence in plant cells.

39. The expression cassette of claim 38, wherein said nucleotide sequence hybridizes to the nucleotide sequence of SEQ ID NO: 1.

40. The Expression cassette of claim 38, wherein said regulator nucleotide sequence is a cauliflower mosaic virus promoter.

41. A method of feeding animals, comprising:

providing a feed comprising matter derived from the transformed plant of claim 17 to an animal for consumption by the animal.

FIG. 1

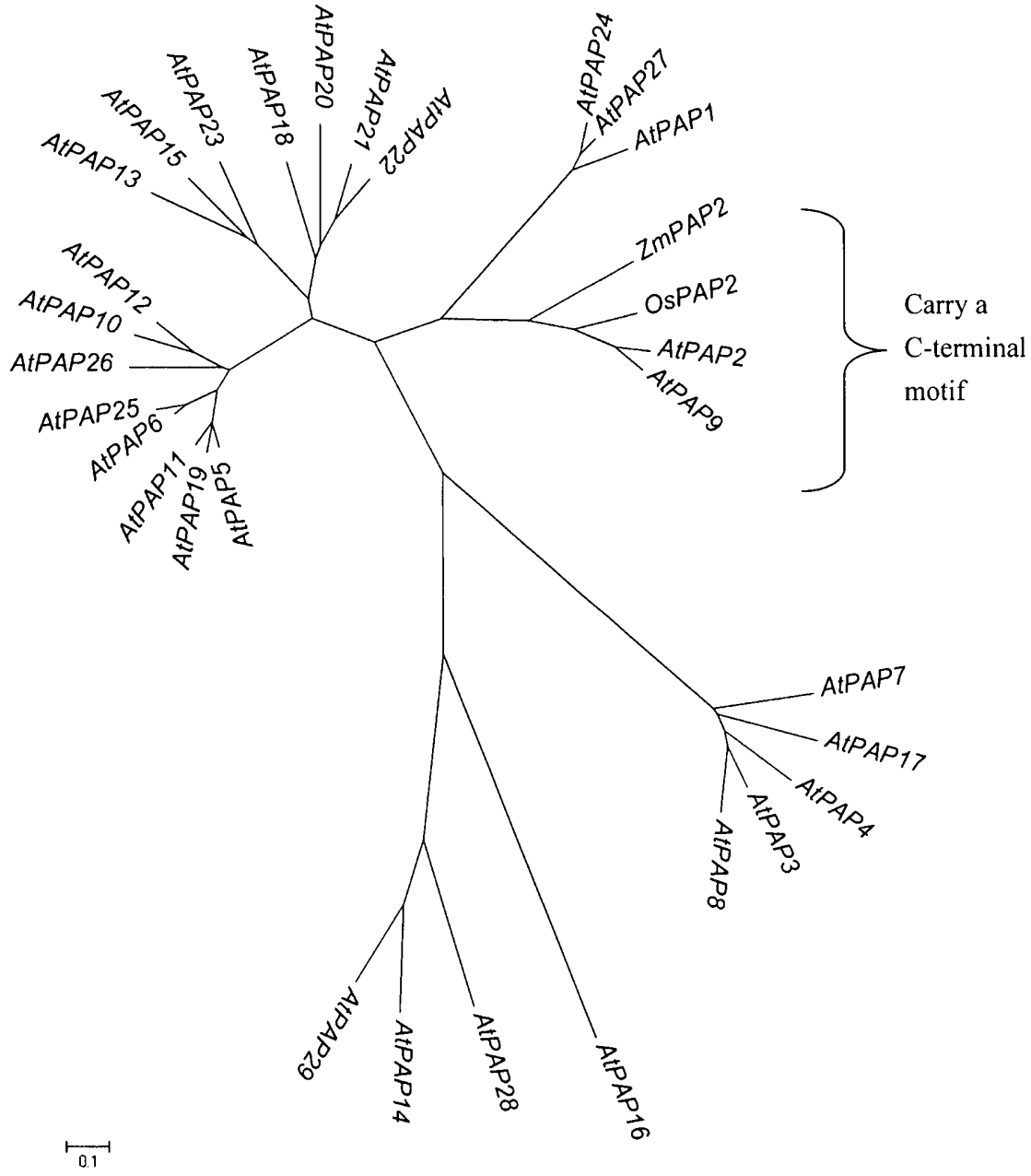

AtPAP20
AtPAP21
AtPAP22
AtPAP18
AtPAP23
AtPAP15
AtPAP13
AtPAP24
AtPAP27
AtPAP1
ZmPAP2
OsPAP2
AtPAP2
AtPAP9
AtPAP12
AtPAP10
AtPAP26
AtPAP25
AtPAP6
AtPAP11
AtPAP19
AtPAP5
Carry a
C-terminal
motif
AtPAP7
AtPAP17
AtPAP4
AtPAP3
AtPAP8
AtPAP29
AtPAP14
AtPAP28
AtPAP16
0.1

FIG. 2A

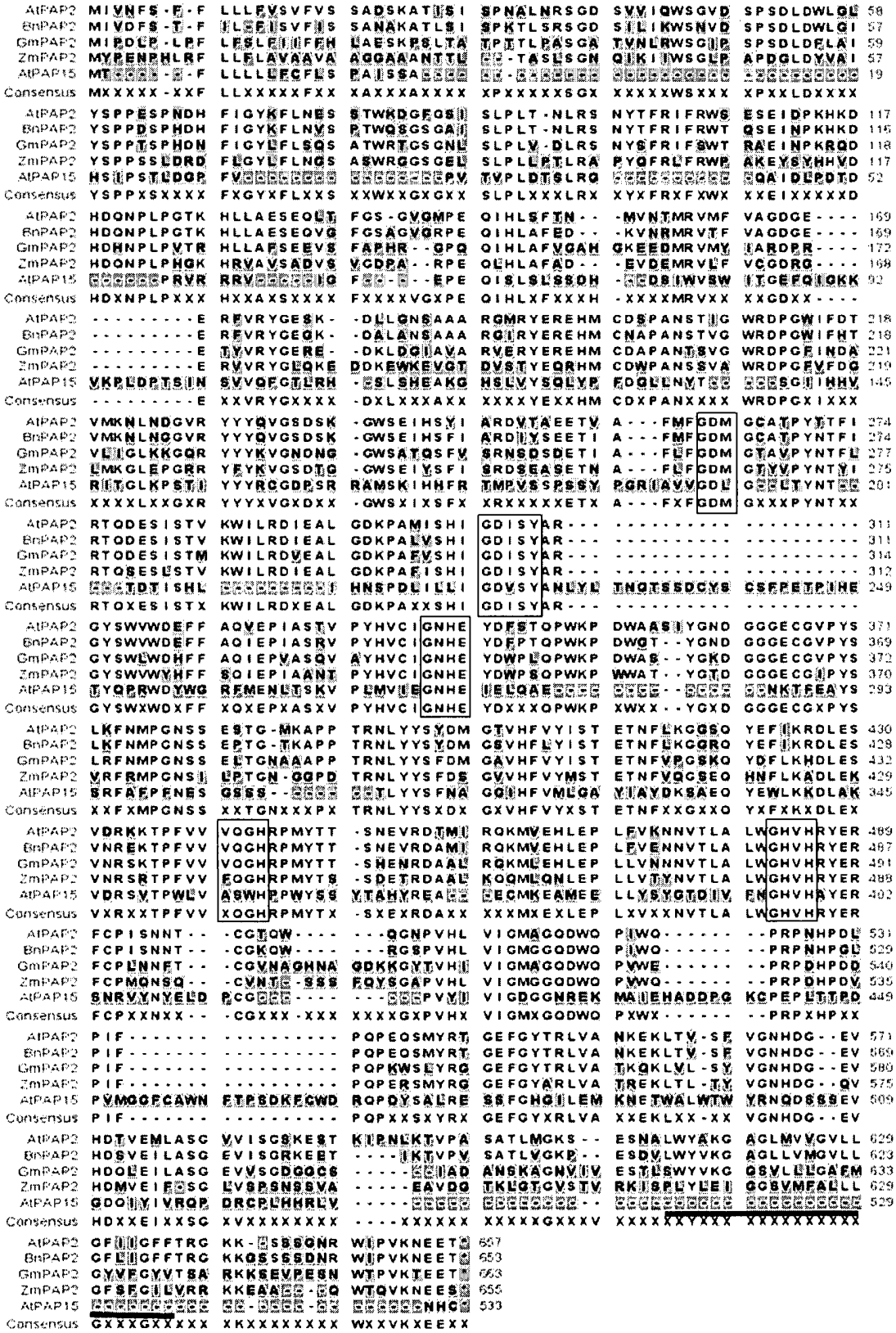


FIG. 2B

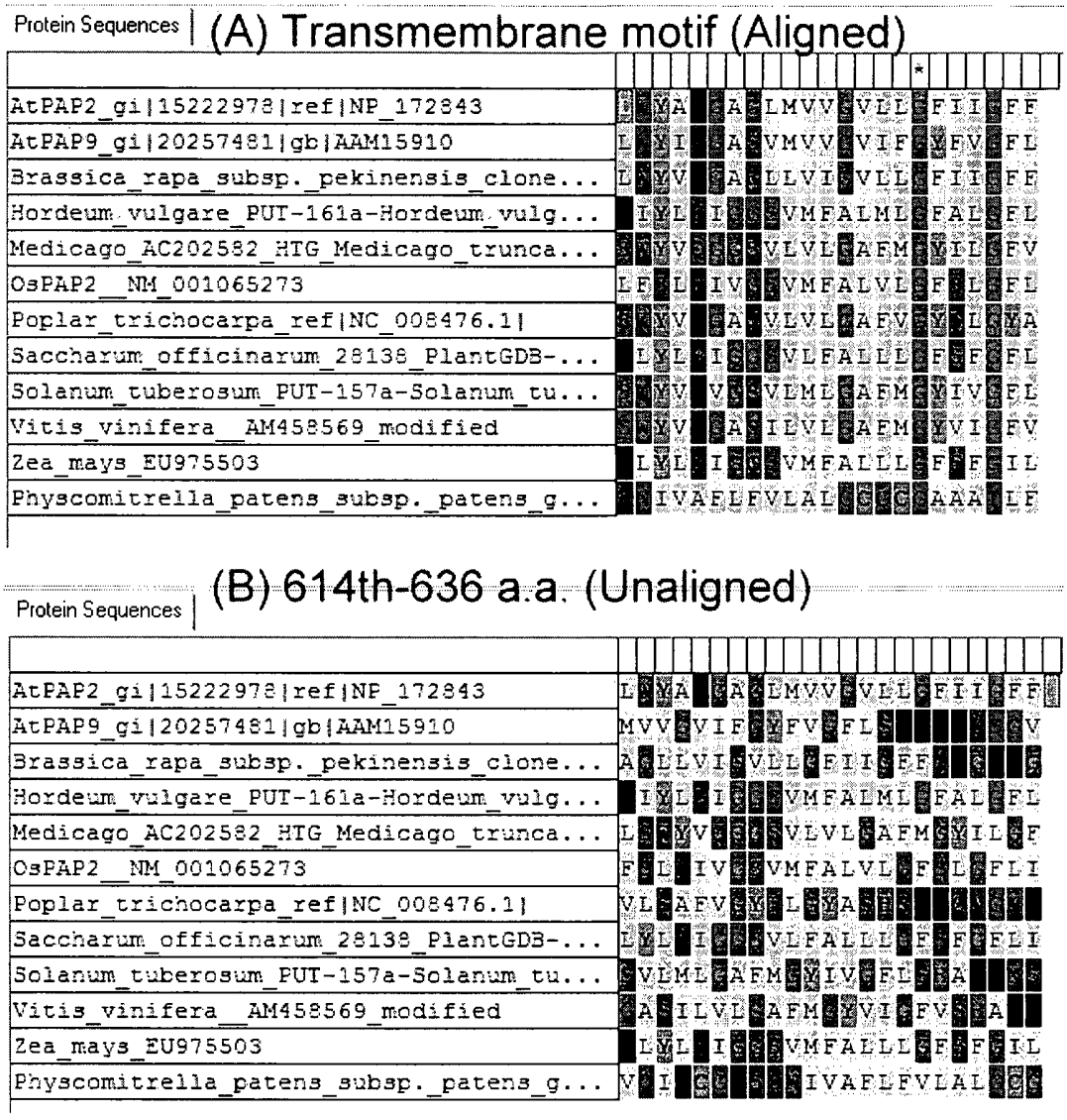

Protein Sequences
(A) Transmembrane motif (Aligned)
AtPAP2_gi|15222978|ref|NP_172843
AtPAP9_gi|20257481|gb|AAM15910
Brassica_rapa_subsp._pekinensis_clone...
Hordeum_vulgare_PUT-161a-Hordeum_vulg...
Medicago_AC202582_HTG_Medicago_trunca...
OsPAP2__NM_001065273
Poplar_trichocarpa_ref|NC_008476.1|
Saccharum_officinarum_28138_PlantGDB-...
Solanum_tuberosum_PUT-157a-Solanum_tu...
Vitis_vinifera__AM458569_modified
Zea_mays_EU975503
Physcomitrella_patens_subsp._patens_g...
Protein Sequences
(B) 614th-636 a.a. (Unaligned)
AtPAP2_gi|15222978|ref|NP_172843
AtPAP9_gi|20257481|gb|AAM15910
Brassica_rapa_subsp._pekinensis_clone...
Hordeum_vulgare_PUT-161a-Hordeum_vulg...
Medicago_AC202582_HTG_Medicago_trunca...
OsPAP2__NM_001065273
Poplar_trichocarpa_ref|NC_008476.1|
Saccharum_officinarum_28138_PlantGDB-...
Solanum_tuberosum_PUT-157a-Solanum_tu...
Vitis_vinifera__AM458569_modified
Zea_mays_EU975503
Physcomitrella_patens_subsp._patens_g...

FIG. 3

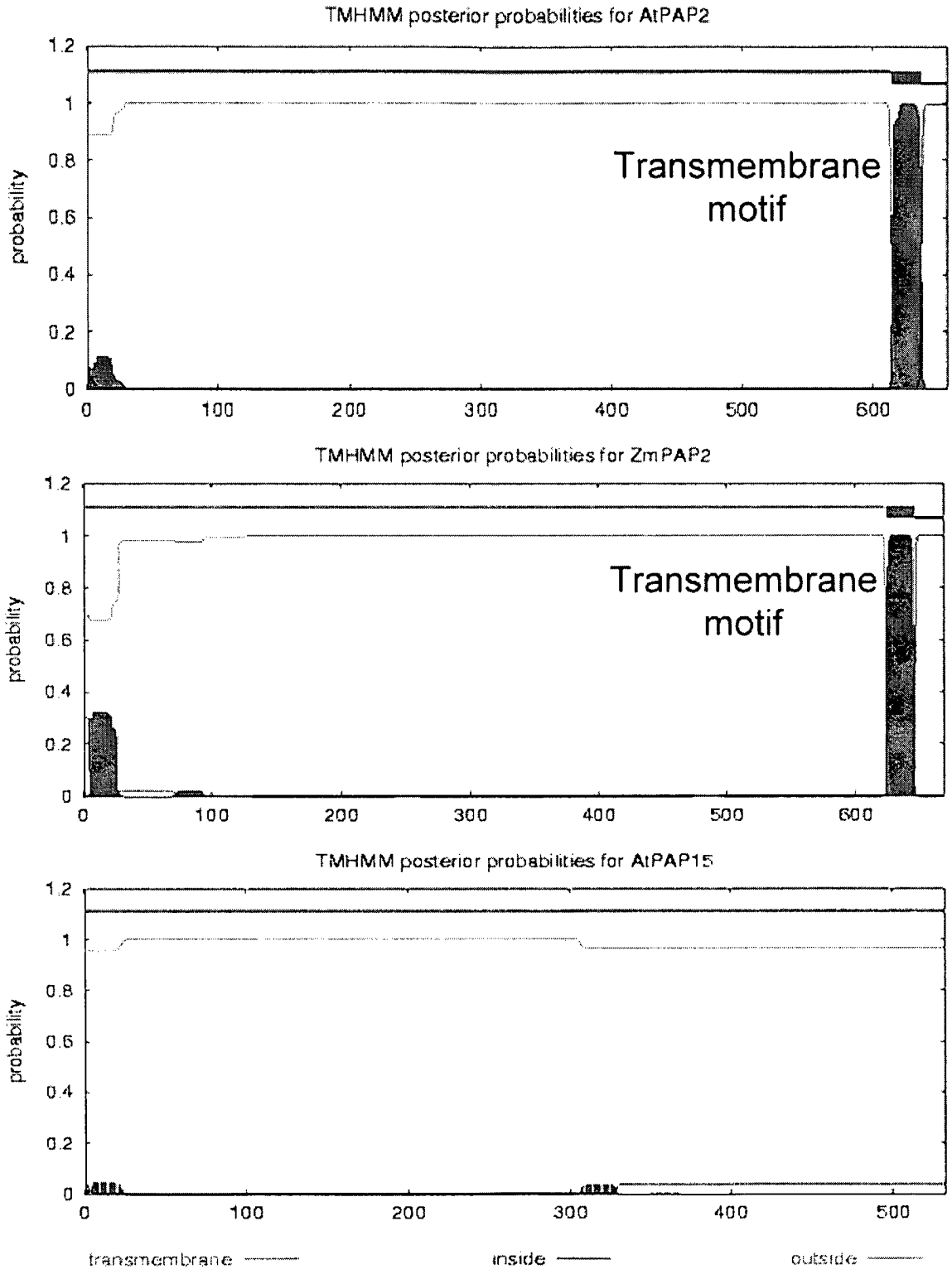
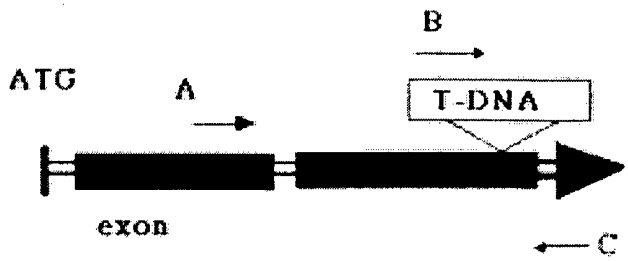

TMHMM posterior probabilities for AtPAP2
probability
Transmembrane motif
TMHMM posterior probabilities for ZmPAP2
probability
Transmembrane motif
TMHMM posterior probabilities for AtPAP15
probability
transmembrane
inside
outside

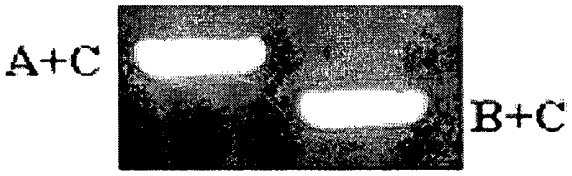
FIG. 4



atpap2-8 structure

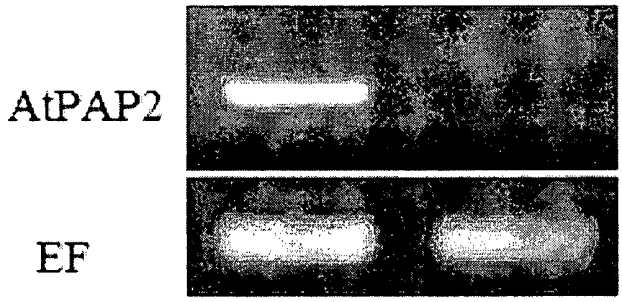
(a) Genomic PCR

WT *atpap2-8*



(b) RT-PCR

WT *atpap2-8*



(c) Western blotting

WT *atpap2-8*

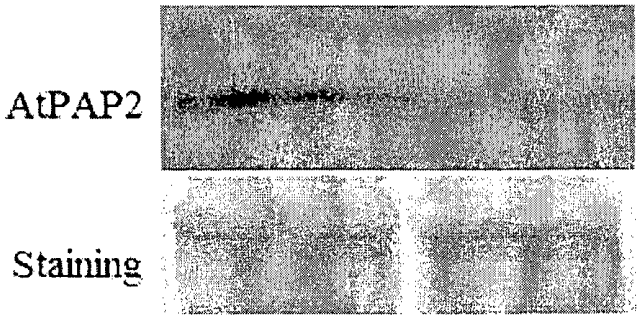


FIG. 5

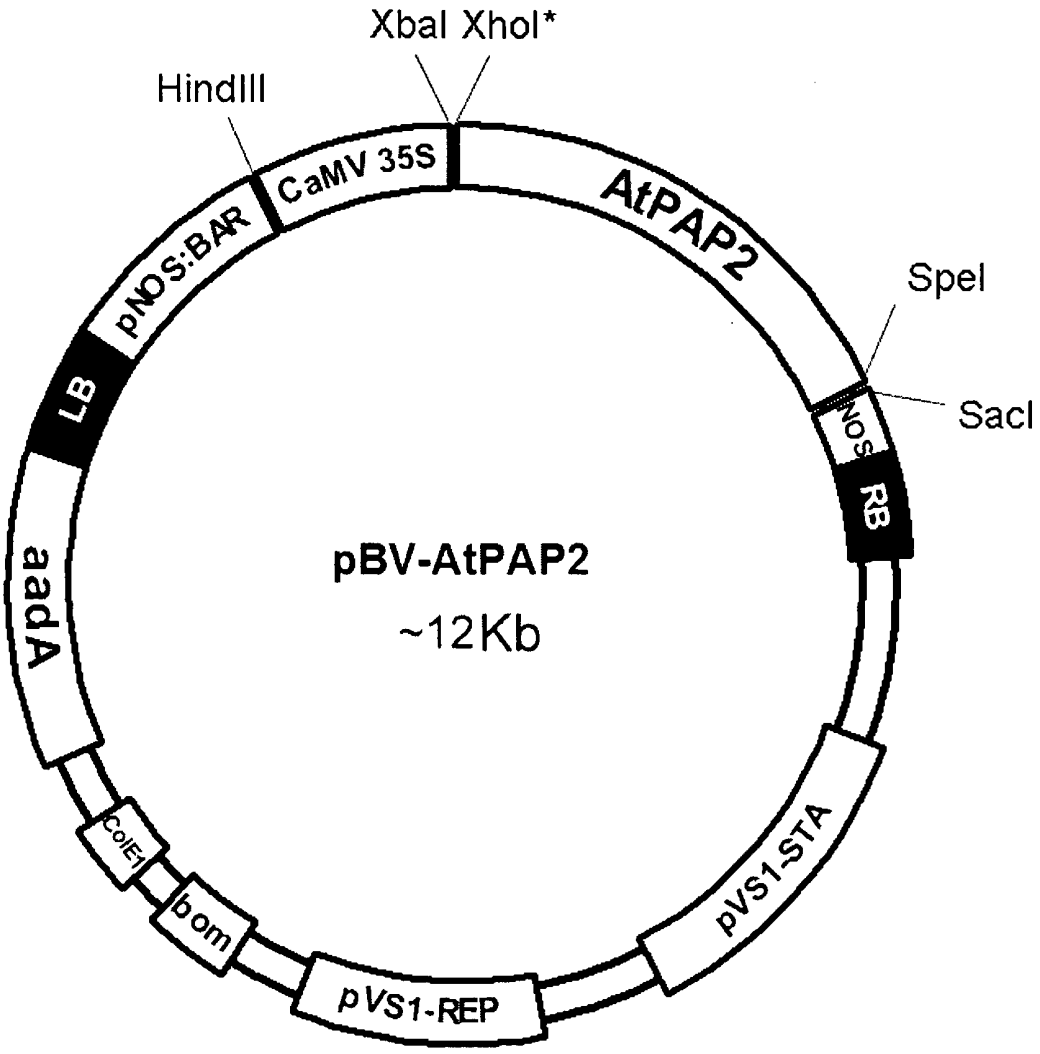

XbaI XhoI*
HindIII
CaMV 35S
AtPAP2
SpeI
SacI
NOS
RB
LB
pNOS:BAR
aadA
ColE1
bom
pVS1-REP
pVS1-STA
pBV-AtPAP2
~12Kb

FIG. 6

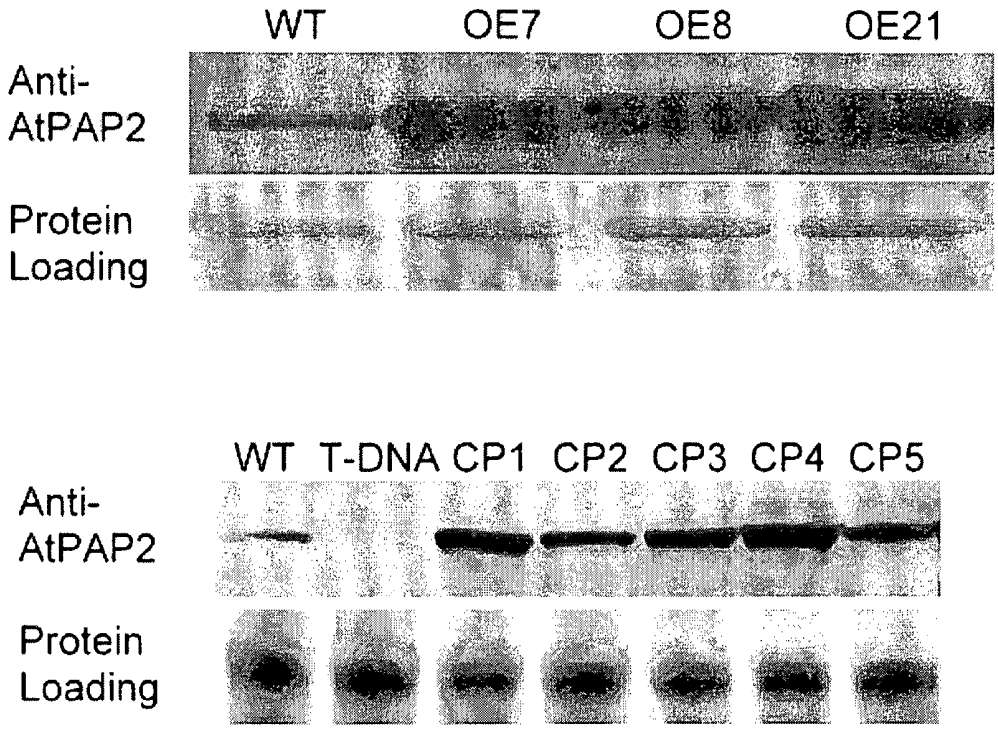
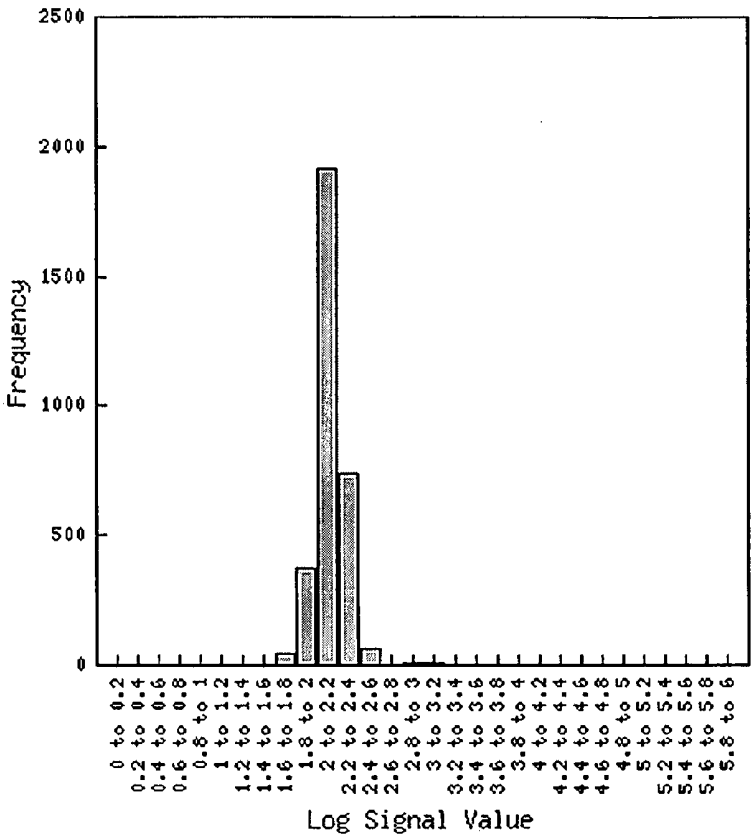
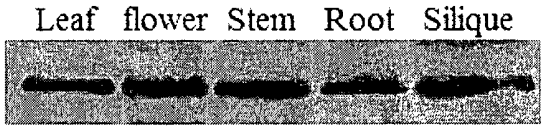

WT
OE7
OE8
OE21
Anti-AtPAP2
Protein Loading
WT
T-DNA
CP1
CP2
CP3
CP4
CP5
Anti-AtPAP2
Protein Loading
B
WT
C15-a
C15-b
C15c
WT
C15-a
C15-b
C15-c
Anti-AtPAP15
Protein Loading
Root
Shoot

FIG. 7

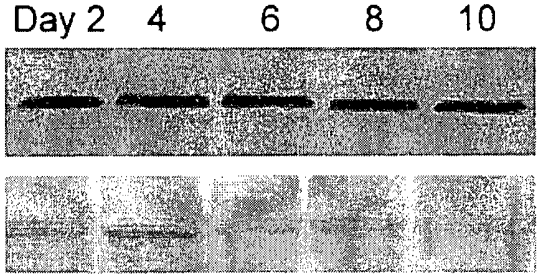
(a)



(b)



(c) Day after germination



(d)

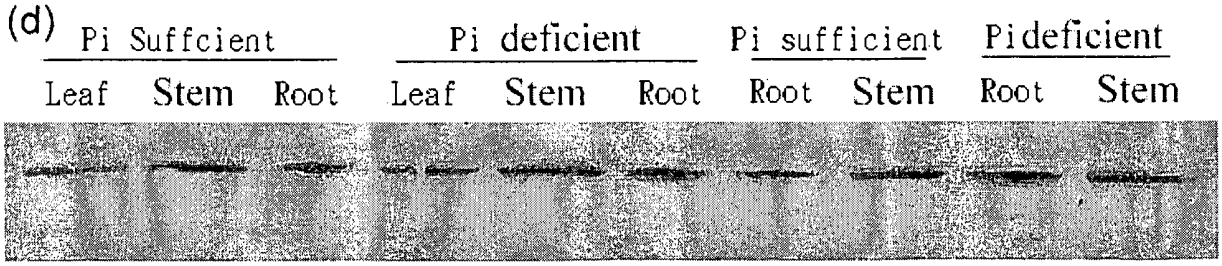


FIG. 8

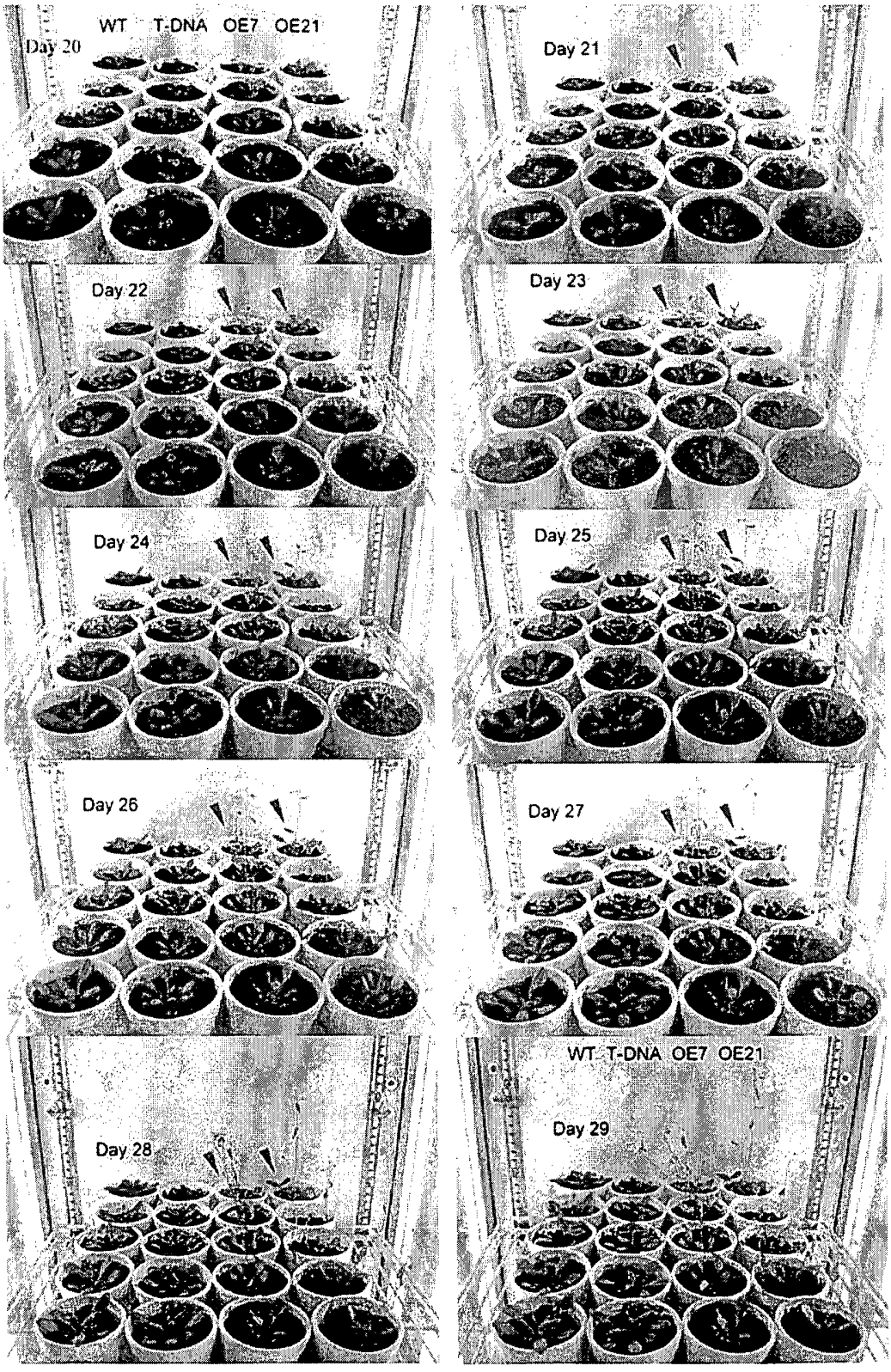
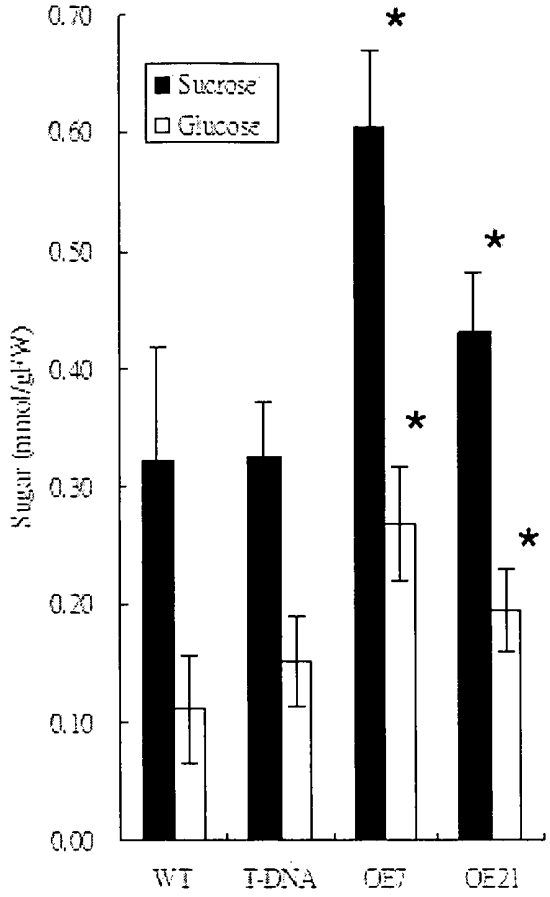

WT T-DNA OE7 OE21
Day 20
Day 21
Day 22
Day 23
Day 24
Day 25
Day 26
Day 27
WT T-DNA OE7 OE21
Day 28
Day 29

FIG. 9



*Statistically ($p<0.001$) different from the WT ($n = 10$).

FIG. 10

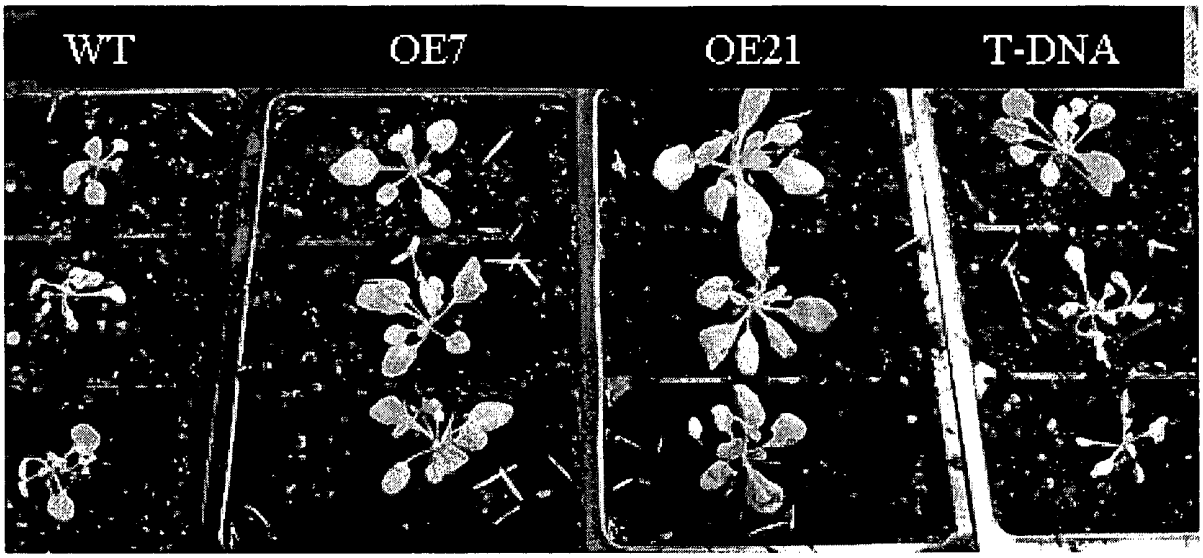


FIG. 11

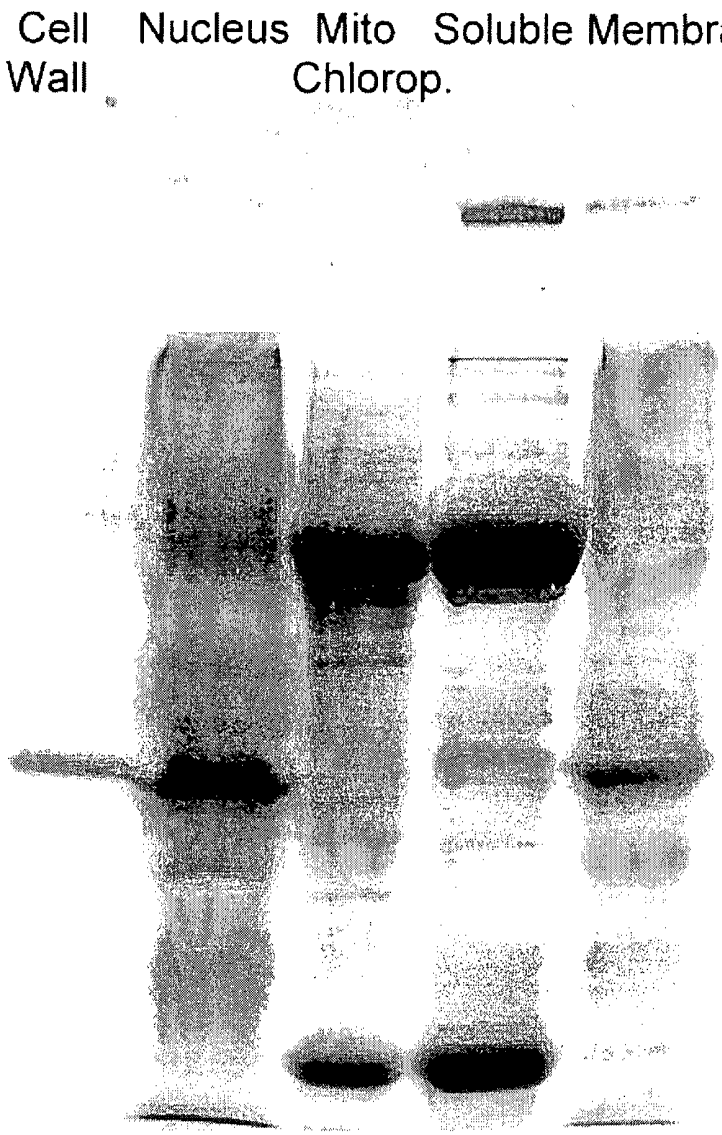


FIG. 12

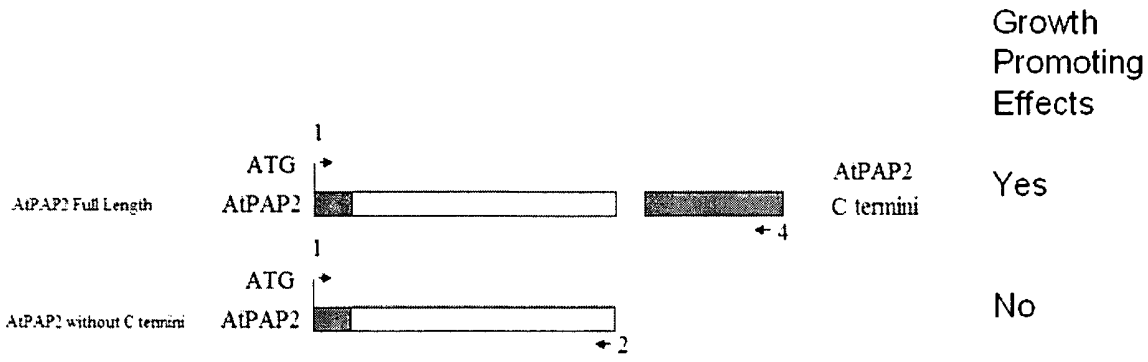
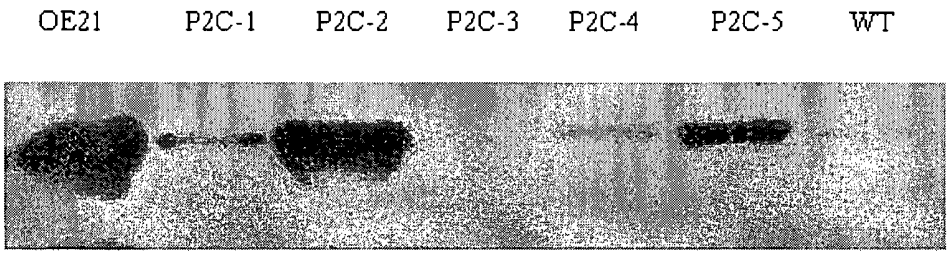

Growth Promoting Effects
1
ATG
AtPAP2
AtPAP2 Full Length
AtPAP2 C termini
4
Yes
1
ATG
AtPAP2
AtPAP2 without C termini
2
No

FIG. 13



INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2009/001465

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: C12N; A01H; A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CPRS; CNKI; WPI; EPODOC; MEDLINE and keywords: purple acid phosphatase, PAP, AtPAP, ZmPAP, OsPAP, transgen+, etc.

EMBL; GenBank: SEQ ID NOs: 1-66.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	XIAO, K. et al. Improved phosphorus acquisition and biomass production in *Arabidopsis* by transgenic expression of a purple acid phosphatase gene from *M. truncatula*. Plant Science. 24 Aug. 2005(24.08.2005), vol. 170, No. 2, pages 191-202, ISSN 0168-9452	
D, X	abstract, page 192, right column, line 21-page 193, left column, line 30, page 194, left column, lines 42-45, Fig. 4	1-4, 13-15, 17-21, 23-24, 26-36, 38-40
D, Y	abstract, page 192, right column, line 21-page 193, left column, line 30, page 194, left column, lines 42-45, Fig. 4	1-41

☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 Feb. 2010(24.02.2010)	**18 Mar. 2010 (18.03.2010)**
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer LIN, Junkai Telephone No. (86-10)62411095

Form PCT/ISA /210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2009/001465

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
D, Y	ZHU, H. F. et al. Expression patterns of purple acid phosphatase genes in *Arabidopsis* organs and functional analysis of *AtPAP23* predominantly transcribed in flower. Plant Molecular Biology. November 2005, vol. 59, No. 4, pages 581-594, ISSN 0167-4412 table 1	1-41
Y	GenBank Accession No. ACG47621, 10 Dec. 2008(10.12.2008), [retrieved on 24 Feb. 2010(24.02.2010)]. Retrieved from: GenBank database. the whole document	1-41
Y	GenBank Accession No. BAC15853, 16 Feb. 2008(16.02.2008), [retrieved on 24 Feb. 2010(24.02.2010)]. Retrieved from: GenBank database. the whole document	1-41
P, X	CN 101475960A (UNIV SOUTH CHINA AGRIC) 08 Jul. 2009(08.07.2009) abstract, examples 1-3 of the description	1-4, 13-15, 17-21, 23-24, 26-36, 38-40
P, Y		1-41
A	WO 2008022570A1 (UNIV CHINESE HONG KONG) 28 Feb. 2008(28.02.2008) the whole document	1-41
A	WO 2008134372A2 (ROBERTS NOBLE FOUND INC SAMMUEL) 06 Nov. 2008(06.11.2008) the whole document	1-41

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2009/001465

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See extra sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2009/001465

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101475960A	08.07.2009	None	
WO 2008022570A1	28.02.2008	US 2009038029A1	05.02.2009
		CN 101541165A	23.09.2009
WO 2008134372A2	06.11.2008	WO 2008134372A3	19.03.2009
		US 2009083874A1	26.03.2009
		AU 2008245794A1	06.11.2008
		EP 2140011A2	06.01.2010

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2009/001465

Continuation of: Box No. III

This Authority considers that there are 3 inventions covered by the claims indicated as follows:

1: claims 1-25, directed to a method to make a transgenic plant comprising introducing a gene coding for a phosphatase into a plant and a transgenic plant made by the said method;

2: claims 26-40, directed to a vector comprising a phosphatase gene, a host cell comprising the said vector, a method for preparing a cell or progeny thereof by using the said vector and an expression cassette comprising a phosphatase gene;

3: claim 41, directed to a method of feeding animals by using the said transgenic plant comprising at least one additional gene coding for a phosphatase.

The reasons for which the inventions are not so linked as to form a single general inventive concept, as required by Rule 13.1 PCT, are as follows: the common or corresponding technical feature of inventions 1-3 is the phosphatase comprising an amino acid sequence of one or more selected from SEQ ID NOS: 2, 6, 8, 19, 21, 23, 25, etc.; however, the prior arts (e.g. ZHU, H. F. et al. Expression patterns of purple acid phosphatase genes in *Arabidopsis* organs and functional analysis of *AtPAP23* predominantly transcribed in flower. Plant Molecular Biology. November 2005, vol. 59, No. 4, pages 581-594, ISSN 0167-4412; GenBank Accession No. ACG47621 and GenBank Accession No. BAC15853) have disclosed the said phosphatase. In conclusion, these claims are not linked by common or corresponding special technical features and define 3 different inventions not linked by a single general inventive concept. The application, hence does not meet the requirements of unity of invention as defined in Rules 13.1 and 13.2 PCT.

CLASSIFICATION OF SUBJECT MATTER

C12N 15/82 (2006.01) i

A01H 5/00 (2006.01) i

A01K 67/00 (2006.01) i

C12N 9/16 (2006.01) i